Latent Factor Analysis of EPSi Conducted in a Mixed General and Clinical Sample

Liz Mutina and Madi Augustine
Khoury College of Computer Science
Northeastern University

Abstract

1. Introduction

Primary care appointments serve as a point of contact for doctors to connect with patients and an opportunity to extend resources and referrals. To effectively coordinate patients with specialty care, quick and efficient methods of screening are essential. A screener must be short enough that it can be done without undue inconvenience at an annual visit, but thorough enough to accurately identify symptoms. As many as 3 in every 100 people will have a psychotic episode in their lifetime[1], and the uniquely rural landscape of Maine serves to benefit from improved methods of detection for early indicators of psychotic spectrum symptoms (PSS). One method of approaching this is to take advantage of annual wellness visits, which in some cases, is the only healthcare interaction individuals have each year. By developing a brief inventory to aid in early detection, universal screening for psychosis can be an important facilitator for successful help-seeking. Accessibility to the appropriate psychosis-specific resources has a direct impact on timely and accurate medical care.

The present methods of screening for psychosis have primarily been developed in populations with concentrated PSS[6]. The Early Psychosis Screener for Internet (EPSi) was created based on cognitive interviews of 14-35 year olds with and without PSS, and has well documented predictive validity. The EPSi is written at a 5th grade reading level and contains 64 questions that serve as a promising core set of items to analyze for patterns of participant responses.

The primary objective of this research is to determine a subset of the 64 item EPSi that represents underlying constructs deduced by variable response patterns to facilitate efficient screening at primary care visits. To accomplish this objective we implement exploratory and confirmatory factor analysis and utilize clustering methods to identify which questions have the highest impact of the individual scoring positive for psychosis. Reducing the number of questions will decrease the amount of time needed to complete the survey and improve the potential for it to be administered during a primary care appointment. This is ideal for rural Maine communities that may have limited access to early intervention.

2. Background

Psychosis involves an altered experience of reality and symptoms can present in a variety of aspects of an individual's life. These symptoms may emerge perceptually, as delusional beliefs, with disordered speech, or in other thoughts and behaviors. Characteristics that are often exhibited by someone experiencing psychosis including auditory or visual hallucinations, cognitive difficulties, paranoia, and disorganized thought patterns[2]. Early detection can reduce the duration of untreated psychosis (DUP) and is associated with better outcomes for those with psychotic spectrum symptoms (PSS). Rural areas often lack the resources to effectively screen individuals for psychosis, hence the overall need to have a short list of potential questions to identify those at high risk for psychotic disorders.

The EPSi was developed from the Early Psychosis Screener (EPS). The development of the EPS began with 490 items during pool development. These items were binned into different aspects of psychosis, and through

cognitive interviewing with subjects reduced to 148 items[3]. Following this initial reduction, machine learning was then used to further reduce the questions using Spectral Clustering Analysis and Support Vector Machine Classifiers. Cross- validation was obtained through having the sample who took both the new 64 question EPSi and Interview for Psychosis-risk Syndromes (SIPS) a previously validated assessment. The EPSi was validated and identified individuals likely to develop psychotic disorder within the next year[4].

This study augments current analyses and aims to examine the latent structure of the Early Psychosis Screen for Internet and reduces the total number of items promoting faster completion of the inventory.

3. Methods/Analysis

For data analysis data was stored via CSV, and PANDAS were used to read in the data and create the dataframe. The sample was collected by Maine Health through the EMES study at primary care clinics and specialty care clinics. The sample size was n=107 (63 female, 43 male,1 prefer not to say). There were 13 participants with active PSS who were engaging in specialty care. The median age for all participants was 17, with ages ranging from 12 to 25 years old.

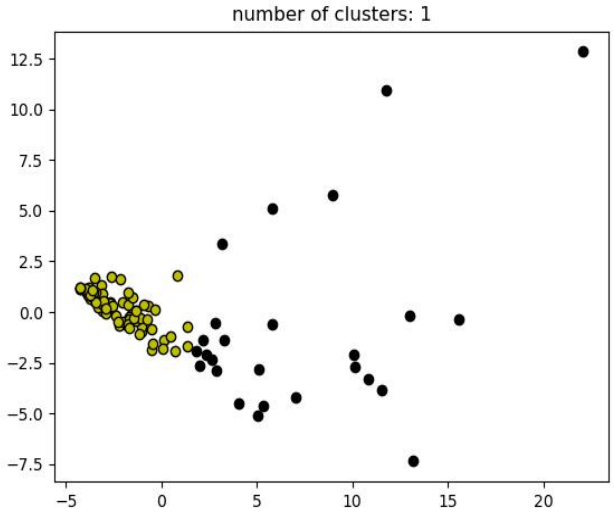
In total this dataset has two timepoints, we utilized the first time point only. Any missing data from timepoint one was replaced with the value that the participant gave in timepoint two. We started with n=107 and removed the 13 PIER participants, therefore the final analysis utilized n=94. PIER participants were individuals who actively engage in a coordinated speciality care clinic and have active PSS.

We used a multiple algorithmic approach to produce the final reduced 16 item set. We aimed to complete factor analysis with the ultimate goal of dimension reduction which was aided by principal component analysis (PCA). To do this we also wanted to see how the data clustered together, so we also utilized fuzzy clustering, gaussian mixture model (GMM), and density-based clustering algorithm(DBSCAN).

4. Results

Initially we wanted to get a visualization of how the 64 variables were clustered together. We began the analysis by using DBSCAN, since it allows for irregular cluster patterns. For graphing, PCA was used to transform the variables to 2D space.

Figure 1. DBSCAN Analysis (n=107)



This data only formed one substantial cluster in yellow and remaining black points were better attributed as noise. Since only one cluster formed the results of this analysis were not significant.

To determine the number of clusters for the following algorithms, a scree plot was utilized.

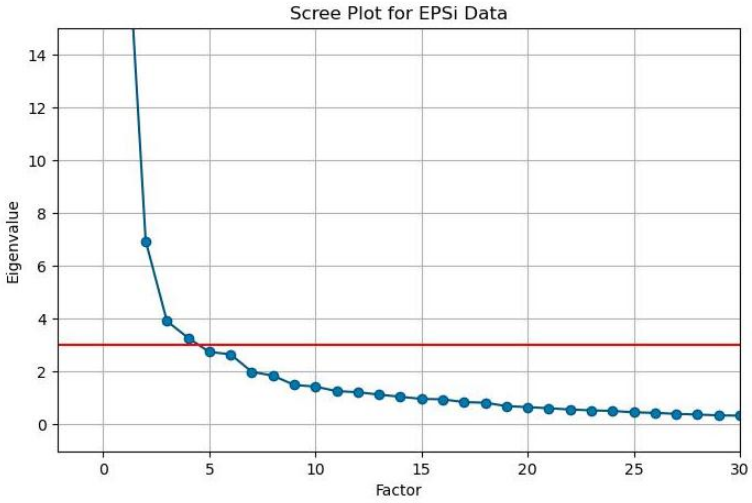


Figure 2. Scree Plot

The Eigenvalue listed on the x-axis was selected to reflect the point of maximum curvature. Based on the Scree plot for this data, a three-factor model was determined as the most optimal for explaining the variance in the data. To validate this interpretation, both two and four-factor models were attempted as well. Ultimately the 2 and 4 factor models were statistically inferior to the 3 factor which contributed to notable under- and overfitting of the data.

We next used fuzzy clustering to see if considering clusters that might overlap aids in the understanding of the data for ultimate factor reduction. Again, PCA was used to transform the data to a 2D space for visualization.

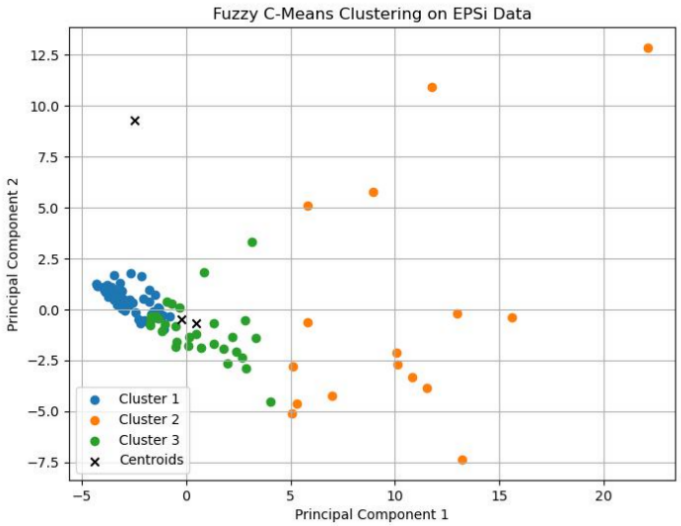


Figure 3. Fuzzy C- Means (n=107)

As visualized in the graph, cluster 2 and 3 had significant overlap, whereas cluster 1 was not closely tied to its centroid and contained more of the outlier variables. Since cluster two and three remained so closely tied together, no meaningful conclusions could be drawn that would ultimately lead to a reduction in EPSi items.

Following observation of the initial clustering patterns, exploratory factory analysis was performed. We initially started with all 64 variables, the Kaiser-Meyer-Olkin(KMO) score was .741 which is middling/moderate or average. The Bartlett test however showed high correlation between variables (p-value = 0.000e+00). Since both tests showed positive results this sample was deemed appropriate for factor analysis.

≥	Items	DoF	DoF Baseline	Chi^2 (p-value)	Chi^2 Baseline	CFI	GFI	AGFI	NFI	TLI	RMSEA	AIC	BIC	LogLik
.4	56	1374	1431	4811.8 (0.0)	8050.66	.48	.40	.38	.40	.46	.159	125.76	414.94	48.12

Table 1. Initial Model Fit Statistics

The initial analysis was attempted with the 3 factor model but showed generally unfavorable statistical fit indices when the model was tested. This was the primary indication that the base dataset contained two distinct populations that do not perform well when integrated for the analyses.

Low-loading variables
(<0.4 on all factors)

12. I thought the people on TV might be talking about me.
22. I thought I had superhuman powers.
26. I thought I was a genius.
56. I needed less sleep than usual.
58. I felt interested in everything.
64. I cared about how other people felt.

Cross-loading variables
(≥0.4 on more than one factor)

3. I thought I was outside my body observing my own life.
11. I thought people on TV said things because they knew I was watching.
16. I felt strange sensations on or under my skin that I could not explain.
18. I thought I was being followed.
28. I felt like thoughts were being placed in my head against my will.
29. My thoughts were being controlled against my will
44. The voice told me what to do.
51. My mind switched between subjects while I was talking.

Figure 4. Low and Cross- Loading Variables

The low loading variables (6) and cross loading variables (8) were removed from the initial variables pool, the subsequent factor analysis was performed using the remaining 50 items.

The results of the GMM model did further inform the poor results in the initial factor analysis.

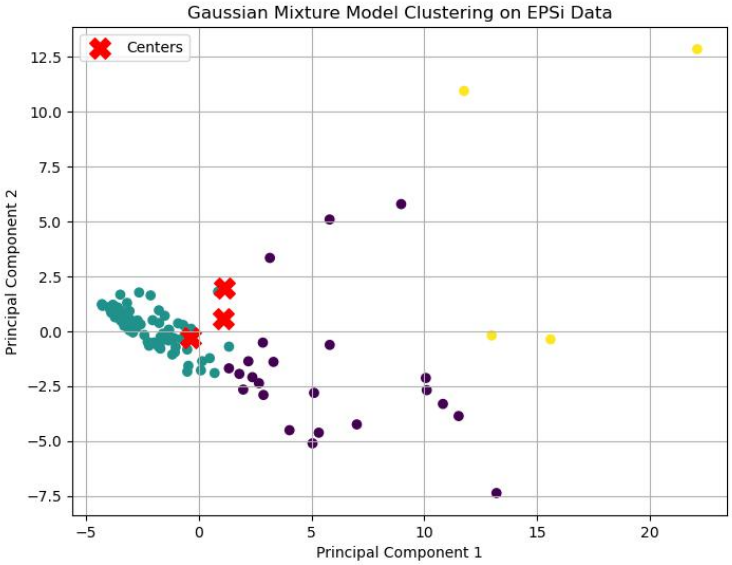


Figure 5. GMM Clustering (n=107)

Our GMM model divided the data into three clusters each with their own distribution. We noticed the yellow cluster only contained only four data points or participants. Review of the data for those four points revealed higher response values to the psychosis screener questions.

```
Average EPSI Score- all variables (Likert 1-5):
  - PIER Program: 2.103
  - Primary Care: 1.510

Average Number of EPSI Items Endorsed (>1):
  - PIER Program: 31.615
  - Primary Care: 17.234
```

Figure 6. PIER vs. Primary Care EPSi Response Patterns

Participants in the PIER program (n=13) were then removed from the data set in an effort to improve the analysis. Some individuals in this sub-population have a clinical history of psychotic spectrum disorders and are actively engaging in specialty care. DBSCAN results show a reduction in noise, though only one cluster forms.

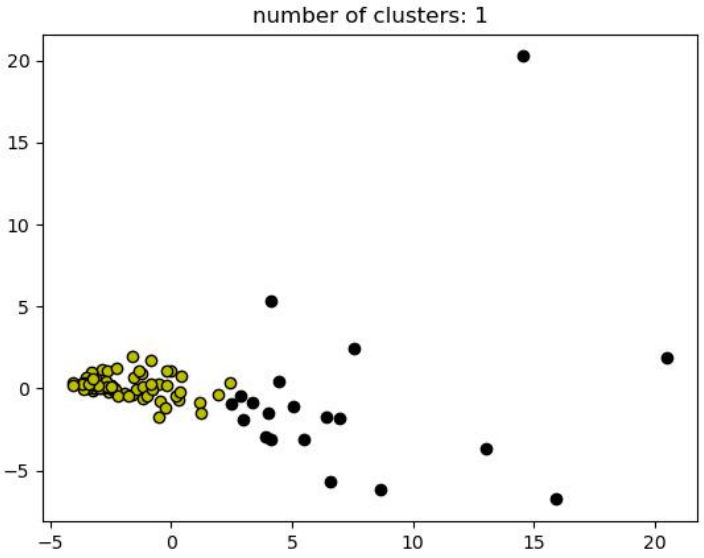


Figure 7. DBSCAN Analysis (n=94)

For the fuzzy clustering model there was an increase in variance between cluster 2 and 3, previously FCP=0.5278 and on the current model is FCP= 0.4734. Cluster 2 (green) displays little to no symptoms of psychosis, these are individuals the screener would not consider further for psychosis intervention.

Cluster 1(orange) is individuals with some mild symptoms of psychosis, they likely would be the group for early identification and potential intervention. They score highest on factors like, "I act without thinking" (0.197), "My mind switched between topics while I was talking"(0.181), "I had so much energy I could not control myself"(0.165), and "I was more talkative than usual" (0.135).

Cluster 3 (blue) would be likely individuals who are experiencing more PSS symptoms and might benefit from more immediate help. They score high on factors like, "I felt like someone was touching me no one was there" (1.023), "I felt like something strange was going on, but I did not know what it was" (1.011), "My mind switched between topics while I was talking" (1.019), "I felt like I had to watch everything around me in order to feel safe"(0.965).

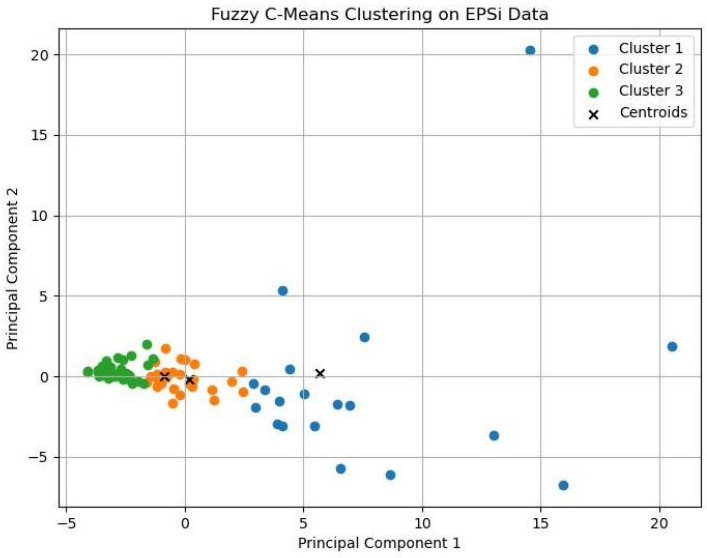
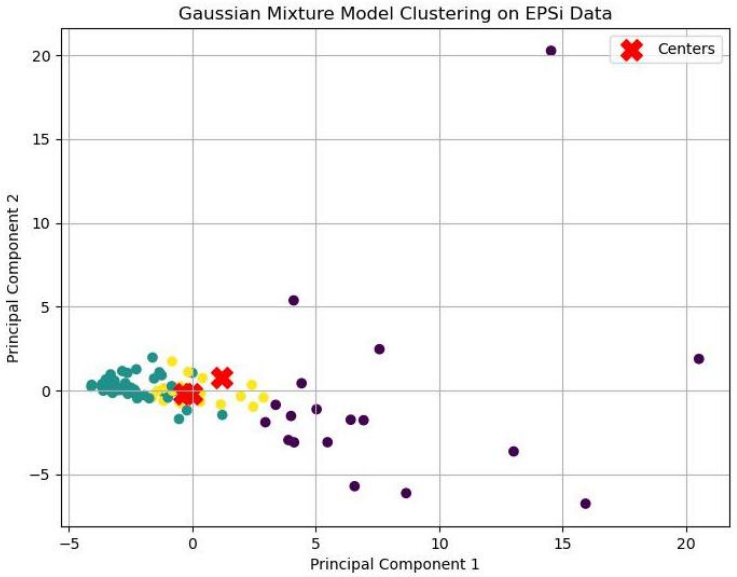


Figure 8. Fuzzy C- Means (n=94)

GMM significantly improved after removing participants from the PIER program. Previously BIC= 12540.69 and AIC= -4656.29, and now BIC= 3106.74 and AIC= -13256.82. This reduction indicates a much better data fit and appropriate model complexity using n=94. Symptom patterns remain similar to fuzzy clustering where there is little to no symptom presentation (teal cluster), moderate presentation (yellow cluster), and higher PSS presentation (purple cluster).

Figure 9. GMM Clustering (n=94)



Similarly to the clustering models, omitting the PIER program participant responses from the exploratory factor analysis yielded fit indices within acceptable range of optimal-fit thresholds. The modified dataset was then confirmed to be suitable for factor analysis (KMO=0.688, and Bartlett's Chi-Square= 7348.03, p-value=0.000e+00). The initial 3 factor model identified low and cross-loading items, and the remaining 50 variables were assigned to the factor that each loaded onto the most. Cross-loading variables were determined by values >0.4 on more than one factor. Questions with loadings <0.4 were dropped from the item set and were considered to be less predictive or specific for the purposes of this analysis.

Below are the 50 EPSi items with associated loading values ≥ 0.4 onto a single factor, and <0.4 onto more than one factor, for the exploratory analysis model:

	Factor1	Factor2	Factor3
epsi_unsure_exp_real	0.449529	0.357143	0.176728
epsi_daydreams_real	0.153951	0.212931	0.487737
epsi_something_strange	0.554309	0.319671	0.305278
epsi_exp_second_time	0.475850	0.116739	0.263871
epsi_hear_my_thoughts	0.269593	0.266466	0.660231
epsi_read_my_mind	0.259432	0.147186	0.688788
epsi_read_others_minds	0.278948	0.228046	0.771705
epsi_messages_things	0.122340	0.023993	0.602759
epsi_special_meaning	0.370263	0.194820	0.406702
epsi_gods_messenger	-0.008626	0.204744	0.860620
epsi_gods_work	0.168623	0.572464	0.071700
epsi_evil	0.248178	0.632357	0.041833
epsi_people_spying	0.232423	0.632319	0.182857
epsi_planning_hurt_me	0.308413	0.664846	0.223730
epsi_plot_against_me	0.382458	0.579333	0.097684
epsi_watching_everything	0.331242	0.551797	0.280552
epsi_predict_future	0.052641	-0.103557	0.692146
epsi_am_famous	-0.045282	0.266273	0.557501
epsi_famous_relationship	-0.113304	0.029717	0.802797
epsi_thoughts_removed	0.381488	0.303417	0.550378
epsi_not_really_exist	0.465938	0.076774	0.137301
epsi_world_real	0.502645	0.149704	0.218120
epsi_famous_romantic	-0.069665	0.017283	0.885408
epsi_romantic_messages	-0.009500	0.013659	0.816538
epsi_sensitive_sounds	0.412833	0.360643	0.180040
epsi_sound_real	0.646047	0.340366	0.038317
epsi_bang_click_hiss	0.524346	0.352088	0.144127
epsi_hear_speaking	0.756012	0.182443	-0.020172
epsi_hear_voices_alone	0.768826	0.139680	-0.081866
epsi_voice_uncertain_real	0.765747	0.281881	-0.067668
epsi_thought_voice_real	0.723473	0.216876	-0.058419
epsi_more_one_voice	0.768661	0.151765	0.100137
epsi_voice_about_me	0.726380	0.228019	0.005930
epsi_voice_mean	0.338845	0.473946	0.020772
epsi_voice_clear	0.759489	0.208663	0.064749
epsi_flashes_flames	0.548598	0.233449	-0.005374
epsi_sensitive_light	0.460865	0.317081	0.247623
epsi_saw_unsure_real	0.688139	0.239694	0.026038
epsi_people_animals	0.758556	0.093117	0.361877
epsi_felt_touching	0.744610	0.173466	0.062141
epsi_energy_trouble	0.319423	0.753542	0.096040
epsi_energy_control	0.283859	0.776843	0.216772
epsi_active_trouble	0.171559	0.633367	-0.019191
epsi_act_without_think	0.184870	0.620704	0.032456
epsi_more_talkative	0.156606	0.584705	0.073559
epsi_many_ideas	0.238177	0.573587	0.136284
epsi_owed_money	-0.074539	0.531712	0.347193
epsi_spent_beyond_means	0.024002	0.527274	0.286856
epsi_bought_expensive	0.214194	0.476305	0.084860
epsi_sexual	0.102526	0.349218	0.611914

Figure 10. 50 EPSi Items with Loading
Each primary loading variable was assigned to the factor with the greatest associated value.

Variance Explained by Each Factor:			
	SS Loadings	Proportion Var	Cumulative Var
Factor1	9.714	0.194	0.194
Factor2	7.609	0.152	0.346
Factor3	7.588	0.152	0.498

Figure 11. Variance Between 3 Factors
For the variance between the factors, Factor 1 is considered the most significant factor (SS= 9.714) since it explains the 19.4% variance in the data. Factor 2 and Factor 3 have similar significance, each explaining 15.2% of the data. Together 49.8% of the data is explained by these three factors.

To further refine the variables, models were developed with all items ≥.4, .5, .6, and .7 respectively, as shown in Table 2 below. Each set of questions that fell at or below the loading thresholds were used to perform confirmatory factor analysis. The factor loading strength is inversely related to the number of questions with the final model represented in the bottom row of Table 2 below. Using a three factor model from these results, we included 16 final items each loading ≥0.7 onto a single factor (<0.4 loading onto a second factor).

≥	Items	DoF	DoF Baseline	Chi² (p-value)	Chi² Baseline	CFI	GFI	AGFI	NFI	TLI	RMSEA	AIC	BIC	LogLik
.4	50	1172	1125	3940.4 (0.0)	6365.42	.46	.38	.35	.38	.43	.159	122.16	384.12	41.92
.5	41	776	802	2666.6 (0.0)	4874.36	.53	.45	.42	.45	.51	.162	113.26	329.44	28.37
.6	28	347	378	1300.3 (0.0)	3047.28	.64	.57	.54	.57	.61	.172	90.33	240.39	13.83
.7	**16**	**101**	**120**	**422.9 (0.0)**	**1680.86**	**.79**	**.75**	**.70**	**.75**	**.76**	**.185**	**61.00**	**150.02**	**4.50**

Table 2. Confirmatory Factor Analysis Fit Statistics

The best threshold performance was at 0.7, this iteration had the highest fit indices (CFI=0.79, GFI=0.75), and lowest AIC/BIC meaning it was the most parsimonious. However, RMSEA (0.185) may indicate potential underfitting of the model, which would warrant further analyses.

Figure 12. Final 16 Items with Factor Loading

(Fig.12) The heatmap representation of the 3 factor model includes all 50 items from the initial exploratory factor analyses. Groupings of high-loading values are highlighted in yellow and bright green, rows with annotated values correspond with EPSi items with ≥ .7 loading values. The associated screener item is written out in full to the left of each variable row.

5. Discussion

Removing EPSi responses of PIER program participants from the final sample yielded significant improvements to all analysis methods. The PIER subgroup tended to score higher on endorsed items compared to the primary care sample and endorsed 50% of the EPSi items on average. By removing these participants, we refine the sample to reflect the target population for the final item set in the context of universal screening. The clustering models confirmed there was a pattern to the responses and displayed how individuals with each degree of symptom severity grouped together. This confirmed the presence of a pattern in responses and symptom severity for respondents that tended to answer questions similarly.

The exploratory factor analysis and confirmatory factor analysis was able to reduce the 64-item inventory to 16 questions.

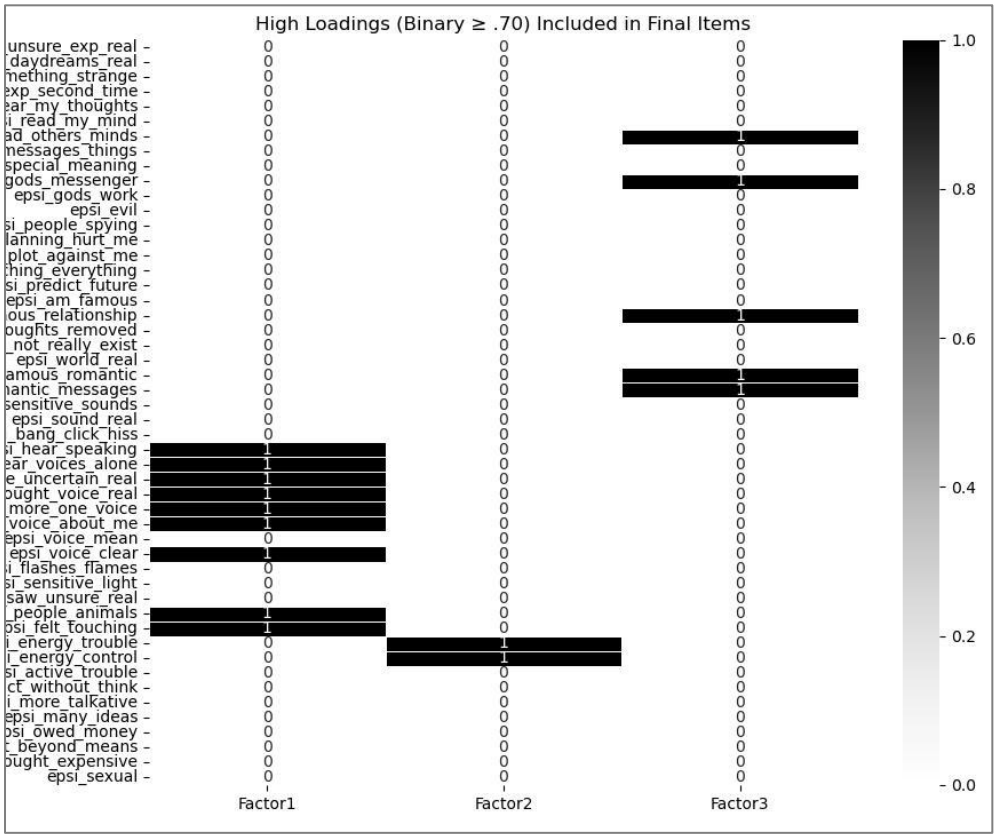


Figure 13. Binary Heatmap of ≥0.7 High Loading Items

We can further explore these results by considering which factors the 16 items loaded onto.

Factor 1 includes questions such as, "*I heard someone speaking to me, but there was no one there.*", "*I thought the voice was real.*", or "*I felt like someone was touching me, but no one was there.*" These questions all go towards potential auditory, visual, or tactile hallucinations, all of which as a section of psychosis based atypical perceptual experiences. Factor 2 included two questions, "*I had so much energy that I got into trouble.*", and "*I had so much energy that I could not control myself*." These experiences of increased activity and self-regulation changes could be indicative of early signs of PSS, which often include shifts in emotions and behavior, as well as disorganization in speech and behavior. Factor 3 included questions such as, "*I thought I could read other people's minds*.", "*I thought I might be God's only messenger on Earth.*", or "*I thought I had a secret romantic relationship with someone famous.*" Grandiose delusions are experiences where an individual has an inflated sense of self-worth, power, or identity. Factor 3 questions are also based on

ideas of reference which is the belief that external events or items have a specific meaning and are able to communicate messages with only the person experiencing them[5].

The conceptual basis for the constructs that direct the underlying factor structure validates the statistical analyses. Cross-loading variables may represent predictive or otherwise important EPSi questions, their endorsement does not present as exclusive to a specific factor. To enhance the underlying constructs each factor represents, it was optimal to select items heavily loaded onto a single factor. This further defines the unique latent constructs that exist naturally within the variable response patterns. Streamlining the variables helped with developing a model that can be interpreted statistically and theoretically. Through this process Factor 1 can be summarized as atypical perceptual experiences, Factor 2 reflects changes to activity level and possible shifts in self-regulation and Factor 3 is weighted mainly by thought disturbances or unusual beliefs. The latent constructs underlying each factor are discrete, discernable, and align with the early symptoms of psychosis.

The final inventory of 16 EPSi items was developed through analysis of 94 participant responses. The participant population for this factor analysis was sampled from primary care offices, making this subset is most suitable for use as a universal screening tool in that setting. Administering this modified 16-item screener is thought to be more feasible and less burdensome, having ¼ the number of questions as the original item pool and requiring less time to complete. The final items were determined to be responsible for the underlying 3 factor constructs within the respondent data, and of particular importance for this sample. By precisely gathering information about certain commonly endorsed symptoms, we may consider use of this instrument in an expanded range of healthcare settings. By identifying elevated symptom burdens for individuals in primary care settings, we can help facilitate help seeking for those at risk of PSS and psychotic disorders.

Limitations

A limitation of this data is the sample contains both symptomatic and non-symptomatic individuals, in future analysis these participants may want to be considered separately to avoid incorrect clustering of the data. The sample itself was also collected from a psychosis-specific specialty care clinic and select primary care offices, a broader sample from the Maine Health system could result in more variability in the responses to the EPSi survey that are not currently present. While we ultimately removed individuals scoring positive for having PSS in the final analysis, the reduction in the sample size may further prevent the data from being representative of the population. Due to the sample's collection at primary care clinics there is also a chance participants undergoing routine care displayed symptoms of PSS and were unaware of symptoms when completing the inventory.

Another limitation is the EPSi survey utilizes a Likert scale, ordinal data represents rank order instead of precise measurements. Participants were not able to express a middle option between 'sometimes' and 'often' for instance. Participants may also interpret the meaning of the terms differently, 'sometimes' for one person may be 'often' to another person.

In the future we would also want to have a full medical and substance use history of individuals completing the inventory. Medical conditions such as Alzheimer's disease, Huntington's disease, dementia, epilepsy, or multiple sclerosis can cause patients to exhibit psychosis-like symptoms during disease

progression. Substances, in particular hallucinogenic drugs, can induce symptoms such as hearing and seeing things that other people cannot. We would want to have a better understanding if the person taking the inventory often uses substances that can elicit psychosis-like symptoms. Understanding the individual's whole history would allow researchers to better control for potential confounding factors.

6. Future work

In the future we suggest first binning the 64 questions into the different areas of psychosis symptomatology. After the questions are broken into the sub-sections factor reduction and clustering could then be completed from there. This would allow the final inventory to be inclusive of all psychosis symptoms someone may display in the early stages of the disease. Without binning the questions, the factors that are reduced might exclude a psychosis symptom area.

We also encourage the research team to revalidate the inventory. One potential first step would be to have individuals take both the newly reduced item inventory and the EPSi inventory and ensure that they score the same level of risk on both assessments. This further investigation would explore if the questions are still predictive of key early psychosis indicators. We would also encourage further exploration in the new order of the questions to evaluate whether there is an influence in scoring based on where the questions are asked in the inventory.

Further, future alternatives for this inventory could utilize a binary yes/no response option instead. It would be worth exploring if the granularity of the Likeart scale adds significant value to diagnosing PSS in the general public during primary care appointments.

As a final reflection the specific aims of the broader EMES study must also be considered. Sensitivity and specificity must be analyzed in both the population with and without PSS. Sensitivity refers to the true positive rate or correctly identifying individuals with PSS, and specificity would refer to the true negative rate, or correctly identifying individuals without PSS. The area under the curve must be evaluated to ultimately assess the predictive validity of the newly reduced questionnaire.

References

1.National Alliance on Mental Illness. (n.d.). *Psychosis*. https://www.nami.org/about-mental-illness/mental-health-conditions/psychosis/

2. Beth Israel Deaconess Medical Center. (n.d.). *Psychosis screening and early intervention: A resource for primary care providers* [PDF]. https://www.psychosisscreening.org/uploads/1/2/3/9/123971055/bidmc_psychosis_pcp_booklet_final.pdf

3. Brodey, B. B., Addington, J., First, M. B., Perkins, D. O., Woods, S. W., Walker, E. F., Walsh, B., Nieri, J. M., Nunn, M. B., Putz, J., & Brodey, I. S. (2018). The Early Psychosis Screener (EPS): Item development and qualitative validation. *Schizophrenia Research*, *197*, 504–508. https://doi.org/10.1016/j.schres.2017.11.027

4. Brodey, B. B., Girgis, R. R., Favorov, O. V., Bearden, C. E., Woods, S. W., Addington, J., Perkins, D. O., Walker, E. F., Cornblatt, B. A., Brucato, G., Purcell, S. E., Brodey, I. S., & Cadenhead, K. S. (2019). The Early Psychosis Screener for Internet (EPSI)-SR: Predicting 12 month psychotic conversion using machine learning. *Schizophrenia Research*, *208*, 390–396. https://doi.org/10.1016/j.schres.2019.01.015

5. Center for Early Detection, Assessment, and Response to Risk (CEDAR), Massachusetts Child Psychiatry Access Project, BCHI / BIDMC. (2019, July). *Psychosis screening in primary care:* https://www.psychosisscreening.org/uploads/1/2/3/9/123971055/bidmc_psychosis_pcp_booklet_final.pdf
6. University of North Carolina School of Medicine, Department of Psychiatry. (n.d.). *Psychosis screening tools*. https://www.med.unc.edu/psych/epi-nc/scope-nc/psychosis-screening-tools/

Acknowledgments

Thank you to Dr. Merelise Ametti and Dr. Erika Mayer at MHIR, Eb Bernier, Research Coordinator and Dr. Kristen Woodberry for allowing us to further explore the EMES data.