4.5 Test retest reliability

> Overall, the analysis demonstrates that some items are more reliable, or perhaps less influenced by change, than others.

Two approaches to retest reliability were conducted for this survey. Firstly, for the first week of fieldwork, community mental health service users who completed a survey were asked if they wished to participate in a retest survey to measure the reliability of the survey. Participants were given one week to return the second survey.

Twenty people agreed to participate in the retest survey and 10 people actually returned a completed survey, giving a response rate of 50%.

With just ten completed returns the sample size was too small to provide a real estimate of the reliability of the survey. However, based on the few surveys available, the surveys show strong correlations between test and re-test for Q5 (.86), Q6 (.76), Q7 (.85), Q9 (.74), Q13 (.65), Q15 (.81), Q19 (.745), Q24 (.85), Q28 (.74) at p ≤ 0.1.

Secondly, in an effort to provide a larger sample of consumers ethical approval was gained to conduct a mail survey with consumers from a community mental health service. To protect consumers' privacy, the service mailed out 217 invitations and consent forms to eligible consumers. Consumers wishing to participate in the study then completed the consent form and received the first survey from Ipsos. Once the first survey was received a second survey was posted to consumers. Consumers received a small payment in recognition of their time and effort in completing the survey. The population response rate through to the second survey was 9% (Table 9).

Table 9: Retest survey response rates

Item	n= 217
Eligible population (A)	217
Consent Forms completed (B)	36
Survey 1 completions (C)	29
Survey 2 completions (D)	19
Opt-in rate (A/B)	17%
Survey 1 response rate (A/C)	13%
Survey 2 response rate (A/D)	9%

As the analysis demonstrated that both scales used have internal properties, Pearson correlation was to measure reliability.

The second survey included a question to identify if consumers had any experiences since they completed the first survey that might have influenced their responses to survey two. Nearly two thirds (62%) had experienced one or more of the following:

- 55% Had contact with someone from their mental health service
- 45% Had contact with another health professional
- 28% Had their medication or treatment changed
- 31% Had a life change (such as housing, employment, finances, relationship, leisure, etc)

Given the significant number of respondents who reported an event between the two surveys, and the potential impact of this on the correlation observed between the test and re-test surveys, moderate levels of correlations were accepted.

Of the 30 questions, 15 were found to be correlated using Pearson's correlation.

Table 10: Correlations between items on the test and retest surveys

Question	Correlation r>.5
Q1	NS
Q2	.445
Q3	.149
Q4	.553
Q5	.729
Q6	.717
Q7	.708
Q8	.283
Q9	.621
Q10	.439
Q11	NS
Q12	NS
Q13	.549
Q14	.543
Q15	.619
Q16	NS
Q17	NS
Q18	.408
Q19	.684
Q20	.580
Q21	NS
Q22	NS
Q23	.739
Q24	.612
Q25	.489
Q26	.623
Q27	NS
Q28	.636
Q29	.584
Q30	.567

NS = not significant
Green shading = correlation >.51

Of the 30 questions, using Pearson's correlation:

- Eight did not have correlations significantly different from zero

- Sixteen had correlations that met the threshold for reliability (r >.51 for moderate correlation)

- Six questions did not have a minimum level of reliability

The seven questions with lower levels of reliability included many topics that might have been influenced by the events consumers reported occurring between completion of the test and retest surveys, such as recontact with the service. It is likely that the variability in results reflects changes in consumers' experience. It is noteworthy that the outcome questions, which measure longer term concepts, were more likely to be reliable than the experience questions.

 Statistical testing of the correlation values (using z' transformations) at 95% confidence level showed that repeated test-retest reliability studies of 11 of the 30 statements would hardly ever return a strong correlation result (r > .70) (see Data Appendix A3).

As there are not strong levels of correlation for all statements, this will affect the ability of those statements to detect a change in consumers' perceived experience of care in longitudinal surveys. The minimum detectable change at 95% confidence level (MDC 95) can be used to highlight differences the statements' ability to detect change in service performance.

4.6 Internal consistency

In constructing the initial survey, to ensure coverage of issues questions were mapped to eight experience domains identified through a review of policy and initial consultations with consumers, carers and professional stakeholders in mental health (Table 11).

Table 11: Policy map of experience questions

Question	Individuality	Choice and involvement	Attitudes, rights, respect	Information	Partnerships	Access	Safety	Physical environment
1. You had opportunities for your family and carers to be involved in your treatment and care if you wanted		●	●		⊙			
2. Your opinions about the involvement of family or friends in your care were respected	●	⊙	●		●			
3. You felt safe to ask questions, provide feedback or make a complaint if you wanted		●	●	●			⊙	
4. Staff made an effort to see you when you wanted		●	⊙			●		
5. You were able to get in contact with this service when you needed	●			●	●	⊙		
6. You had access to your treating doctor or psychiatrist when you needed		●		●	●	⊙		
7. You had access to a range of other professional services if you needed (such as dietary advice, talking therapies, skill development, etc)	●	●			●	⊙		
8. You felt welcome at this service	●		⊙				●	
9. The facilities and environment met your needs (such as cleanliness, private space, reception area, furniture, common areas, etc)	●					●	●	⊙
10. You were able to do the things that were important to you while using this service (such as have family and friends visit, make phone calls, have a cup of tea or coffee, etc)	⊙	●	●					●
11. Staff caring for you took the time to get to know you as a person	⊙	●	●					
12. Your individuality and values were respected (such as your culture, faith or gender identity, etc)	⊙	●	●					●
13. You were listened to in all aspects of your care and treatment	●	⊙	●					
14. You were involved in planning your future care	●	⊙			●	●		
15. You had opportunities to discuss your progress with the staff caring for you	●	⊙	●					
16. Staff showed respect for how you were feeling	●		⊙					
17. Staff worked as a team in your care and treatment (for example, you got consistent information and didn't have to repeat yourself to different staff)		●	●		⊙	●		
18. Staff ensured you understood the effects of your treatment options (including any medication, talking therapies, etc)		●	●	⊙			●	
19. You felt safe using this service	●		●	●	●	●	⊙	●
20. Your privacy was respected	●	●	⊙		●			●
21. Staff showed hopefulness for your future			⊙					
22. You had things to do that were meaningful for you	⊙	●	●		●			
23. Access to peer support (such as information about peer workers, referral to consumer programs, advocates, etc)		●		●	●	⊙		
24. Convenience of the location for you (such as close to family and friends, transport, parking, community services you use, etc)	●				●	⊙		●
25. Explanation of your rights and responsibilities		●	⊙	●			●	
26. Information given to you about this service (such as how the service works, which staff will be working with you, how to make a complaint, etc)		●	●	⊙			●	●
27. Development of a care plan with you that considered all of your needs (such as health, living situation, age, etc)	⊙	●			●	●		

Internal consistency for each of the eight experience domains was evaluated using Cronbach Alpha.
All alpha values but one (partnerships) demonstrated good internal consistency (Table 12).

Table 11 Key
⊙   Primary domain
●   Secondary domain

Table 12: Alpha values for experience domains

Experience domain	Alpha value
Individuality	.81
Choice and involvement	.85
Attitudes, rights and responsibilities	.89
Information	.72
Partnerships	.44
Access	.80
Safety	.75
Physical environment	n.a. (single question)

4.7 Construct validity

On the full sample, five domains were identified that were more internally consistent than the eight experience domains initially developed. The community and inpatient PCAs explain more variance than the PCA for the full sample. However, the difference is marginal and the models very similar. Across the PCA analyses, access, information and individuality are common themes. Privacy and safety are also commonly linked. The construct domains should be further explored in the Pilot to determine the implications of different constructs for reporting (to services and government).

Construct validity requires the presence of both convergent and discriminant validity. Examining discriminant validity between any two rating questions relies on bivariate correlations as well as the reliability estimate of each question. Hence, given that the re-test did not yield a sufficient number of surveys, discriminant validity cannot be reported. This section will focus on convergent validity which measures whether constructs that should theoretically be related actually are related.

4.7.1 PCA full sample

To examine convergent validity (identifying the eight designated domains from the pattern of ratings), a principle component analysis (PCA) was conducted on the entire data set.

The initial PCA provided two metrics indicating that the set of ratings is suitable for structure detection:

- KMO measure of sampling adequacy: .96
- Bartlett test of sphericity (Chi2 = 3487, 231 df, $p < 0.001$)

The eight domains do not emerge as factors from the data and only three factors emerge (accounting for 68% of total variance) (Table 13).

Table 13: PCS eigenvalues (full sample)

<div align="center">

Initial Eigenvalues

</div>

Domains	Total	% of Variance	Cumulative %
1	12.375	56.252	56.252
2	1.514	6.880	63.131
3	1.034	4.700	67.831

These three domains related to: individuality, information, access (Data Appendix A4). Domain 1 had 13 items. These were subject to a further PCA which yielded two factors. So the final factor analysis on the full sample projected five domains. These domains were found to have a strong relationship with the theorized eight experience domains (Table 14).

Table 14: Comparison of experience and PCS domains (full sample)

Experience domain	PCA domain
Individuality	Respect and listening
Choice and involvement	Respect and listening
Attitudes, rights and responsibilities	Respect and listening
Information	Information
Partnerships	-
Access	Access
Safety	Privacy and safety
Physical environment	Facilities and the environment

It is not surprising that no domain was found to relate to the concept of partnerships as the Cronbach Alpha demonstrated poor internal consistency for these questions (Section 1.10).

The internal consistency analysis repeated on the five PCS domains identified delivers more consistent scores than the eight experience domains (Table 15).

Table 15: Internal consistency of PCS domains (full sample)

Domain	Alpha value
Respect and listening	.93
Privacy and safety	.90
Information	.88
Access facilities and environment	.86 n.a. (single question)

4.7.2 PCA inpatient sample

To determine if the overall model is consistent for the different populations of consumers (inpatient and community), PCAs were conducted on each subgroup.

The initial inpatient PCA provided two metrics indicating that the set of ratings is suitable for structure detection:

- KMO measure of sampling adequacy: .94
- Bartlett test of sphericity (Chi2 = 2598, 351 df, p < 0.001.

Again, the eight experience domains do not emerge as factors. Four factors were found to account for 71% of the variance (Table 16). The solution was rotated to maximise differences in correlation across factors (Data Appendix A6)

Table 16: PCS eigenvalues (inpatients)

Initial Eigenvalues

Domains	Total	% of Variance	Cumulative %
1	15.312	56.713	56.713
2	1.694	6.272	62.985
3	1.298	4.808	67.793
4	.958	3.549	71.342

A PCA was conducted on the first factor, with 14 items, to produce two domains. The final five domains were:

- Individuality
- Privacy and safety
- Information
- Access
- Meaningful activities.

4.7.3 PCA community sample

The PCA was conducted on the sub-set of community patients (but without Q22 as this was only asked of inpatients).

The initial PCA provided two metrics indicating that the set of ratings is suitable for structure detection:

- KMO measure of sampling adequacy: .91
- Bartlett test of sphericity (Chi2 = 1671, 325 df, p < 0.001.

The designated eight domains do not emerge as separate factors, five domains emerge (Table 17).

Table 17: PCS eigenvalues (community)

<p align="center">**Initial Eigenvalues**</p>

Domains	Total	% of Variance	Cumulative %
1	12.231	47.042	47.042
2	1.698	6.531	53.573
3	1.434	5.514	59.087
4	1.320	5.075	64.162
5	1.009	3.879	68.041

The five domains were rotated to maximize differences in correlation across factors (Appendix 6). Factors emerged around the following topics:

- Respect and care
- Access
- Information
- Privacy and safety
- Individuality

4.8 Criterion-related validity

Consumer experience of care is the antecedent to outcome ratings. That is, there is a strong relationship between consumer ratings of care experience and care outcomes. The outcome questions are functioning as intended.

Criterion-related validity analysis examines validity by linking measures external to the survey to survey measures. However, given the anonymous character of the survey, external measures are not available in this survey.

Criterion-related validity analysis was conducted instead by using outcome questions as dependent variables of consumer experience of care:

> *Q28. Overall, how would you rate your experience of care with this service in the last 3 months?*

> *Q29. The effect the service had on your ability to manage your day to day life.*

> *Q30. The effect the service had on your hopefulness for the future.*

> *Q31. The effect the service had on your overall well-being.*

The four outcome questions were found to be highly correlated (r>.7)[4]. (Data Appendix A8).

The relationship between the outcome variables and consumer experiences of care was tested on two polar opposites:

- What facets of consumer experience of care separate people with a negative experience (Poor or Fair) vs a relatively positive one (Good).
- What facets of consumer experience of care separate people with a strong positive experience (Very Good or Excellent) vs a relatively positive one (Good).

Binary logit regression was used to identify which facets of consumer experience of care link to positive or negative outcome, based on each outcome variable.

The binary logit re-classified between 70% and 90% of each defined group of consumers (negative experience as Poor or Fair, Positive as Good, and strongly positive as Very Good or Excellent) (Table 18).

Table 18: Binary logit regression models

Negative to positive				Positive to strongly positive			
Q28. Overall experience	Predicted			**Q28. Overall experience**	Predicted		
	Q28 Low side-binary		Percentage Correct		Q28 High side-binary		Percentage Correct
	.00	1.00			.00	1.00	
Low side-binary .00	39	12	76.5	**Q28** High side-binary .00	37	14	72.5
1.00	14	38	73.1	1.00	11	89	89.0
Overall Percentage			74.8	Overall Percentage			83.4
Q29. Ability to manage day-to-day life	Predicted			**Q29. Ability to manage day-to-day life**	Predicted		
	Q29 Low side-binary		Percenta Correct		Q29 High side-binary		Percenta Correct
	.00	1.00			.00	1.00	
Low side-binary .00	49	13	79.0	High side-binary .00	43	19	69.4
1.00	18	50	73.5	1.00	15	58	79.5
Overall Percentage			76.2	Overall Percentage			74.8
Q30. Hopefulness for the future	Predicted			**Q30. Hopefulness for the future**	Predicted		
	Q30 Low side-binary		Percenta Correct		Q30 High side-binary		Percenta Correct
	.00	1.00			.00	1.00	
Low side-binary .00	50	15	76.9	High side-binary .00	39	26	60.0
1.00	15	46	75.4	1.00	18	59	76.6
Overall Percentage			76.2	Overall Percentage			69.0

In relation to the specific questions identified as key drivers in each model (Figure 6) there was very strong symmetry in the drivers of negative and positive experience. This validates the use of a linear model across the entire spectrum of experience (that is, there is a direct and predictable relationship between experience and outcome questions).

---

[4] As Q31 had r>.9 correlation to two other outcome variables it was removed from the analysis.

Figure 6: Summary of drivers

Experience driving <u>overall experience past 3M</u> (Q28)

Driving negative experience

- Q8: You felt welcome at the service
- Q9: The facilities and environment met your needs
- Q23: Access to peer support
- Q25: Explanation of your rights and responsibilities
- Q27:Development of a care plan with you

Driving strong positive experience

- Q8: You felt welcome at the service
- Q21: Staff showed hopefulness for your future
- Q23: Access to peer support
- Q24: Convenience of the location
- Q27:Development of a care plan with you

Experience driving ability to <u>manage day-to-day</u> (Q29)

Driving negative experience

- Q16: Staff showed respect for how you were feeling
- Q17: Staff worked as a team in your care and treatment (for example, you got consistent information and you didn't have to repeat yourself to different staff)
- Q20: Your privacy was respected
- Q23: Access to peer support
- Q24: Convenience of the location

Driving strong positive experience

- Q6: You had access to your treating Dr or psych when you needed
- Q15: Your had opportunities to discuss your progress with the staff caring for you
- Q17: Staff worked as a team in your care and treatment
- Q18: You felt safe using this service
- Q24: Convenience of the location

Experience driving <u>hopefulness for the future</u> (Q30)

Driving negative experience

- Q10: You were able to do the things that were important to you
- Q21: Staff showed hopefulness for your future
- Q23: Access to peer support
- Q26: Information given to you about this service

Driving strong positive experience

- Q9: The facilities and environment met your needs
- Q12: Your individuality and values were respected
- Q13: You were listened to in all aspects of your care and treatment
- Q23: Access to peer support
- Q24: Convenience of the location