

**Analysis of PGF ORS and SRS data 2011-2014.
Report for the Problem Gambling Foundation – July 2015**

This analysis is based on data collected between 2011 and 2015. All clients in the database started after 31/12/2010 and none started after 31/12/2014, although data up to April 2015 from clients who had started in 2014 or earlier is used.

Overall there were 17756 sessions with 17368 *Outcome Rating Scale* (ORS) complete data sets, and 16941 *Session Rating Scale* (SRS) completed data sets. Data sets marked 0,0,0,0 (the four sub-scales of both the ORS and SRS) were counted as genuine, although it probable that much of this data should have been entered as blank (i.e. the test was not done) or treated as spoiled data.

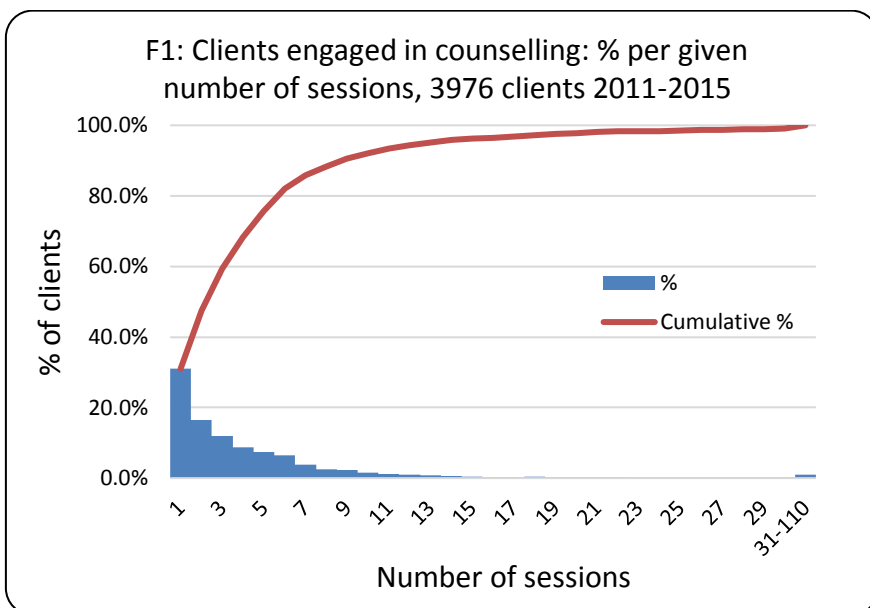


Figure 1 shows that 3976 clients were involved and that the number of sessions each one had varied between 1 and 110, with 31% having just the single session, over half having 3 or more sessions, 92% having 10 sessions or less, and 4.8% having more than 13 sessions.

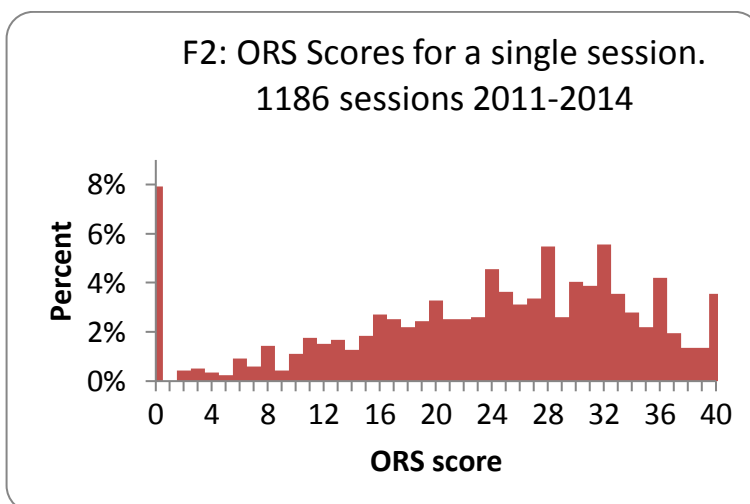


Figure 2 gives a picture of the ORS scores for the 1186 clients who only attended a single session - 96% of all single session data. The ORS self-assessed scale is a very quick and simple measure of current mental health which has been validated against a number of much longer standardised general tests of mental health status. Figure 2 shows a general tendency for self-judgements to be on the positive side of the scale, with a mean of 23, but with 8% rating their mental health as 0. This suggests either these clients are very depressed and disturbed or this is spoiled data possibly indicative of reluctance to attend counselling or a data entry issue.

Figure 2 shows curious pattern in which ORS scores that are a multiple of 4 are generally much more frequent than the surrounding data. This tells that there is substantial number of data sets where the four contributing subscales to the ORS are both whole numbers and likely to have 2-4 numbers which are the same. This means that in many cases the measuring for the contributing scores has been approximate rather than precise. This level of imprecision will not, however, have much effect on the measurements of change that derive from ORS data, but it may limit the sensitivity of the ORS subscales

T1: ORS subscale scores for 1186 single session clients		
Subscales – “How are you feeling – “	mean	SD
individually - personal wellbeing	5.77	2.95
Interpersonally - family, close relationships	5.77	3.01
Socially - Work, school, friendships	5.75	2.99
Overall	5.80	2.90

Table 1 confirms the likelihood that clients respond to the four ORS subscales in a very similar way. The means and standard deviations are almost identical. A one-way ANOVA confirms that there is no significant difference between the subscales ($p=0.70$) and their intercorrelations range from 0.77 to 0.83 – all highly significant ($p=0.000$). This may be an effect of being a first time user of the ORS instrument.

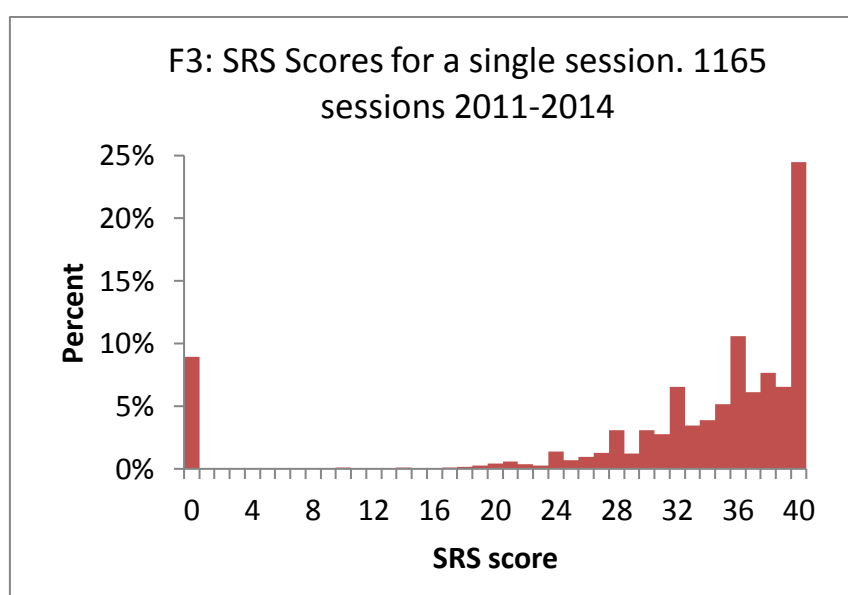
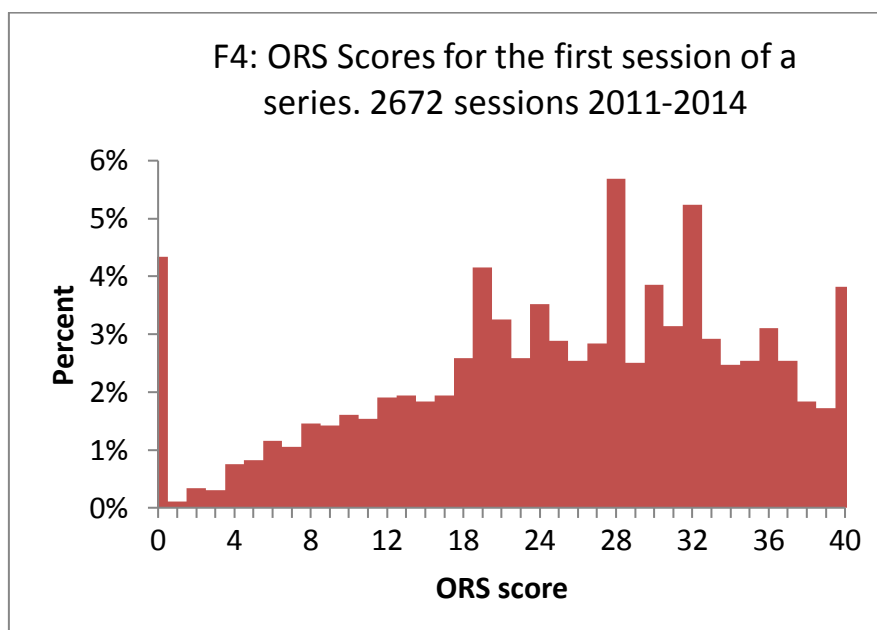


Figure 3 shows that single session clients using the SRS, responded positively with a mean score SRS of 32, 60% scoring over 35 and 80% over 30. The SRS is a measure of client satisfaction with the counselling session and a score 40 (25% scored 40) means that everything in the session went well. On the other hand 9% scored 0, meaning that nothing worked. There is a highly significant correlation of $r=0.48$ ($p=0.000$) between ORS and SRS scores showing people who rated themselves as distressed tended not like the counselling, whereas those who saw themselves as functioning reasonably well, did like counselling.

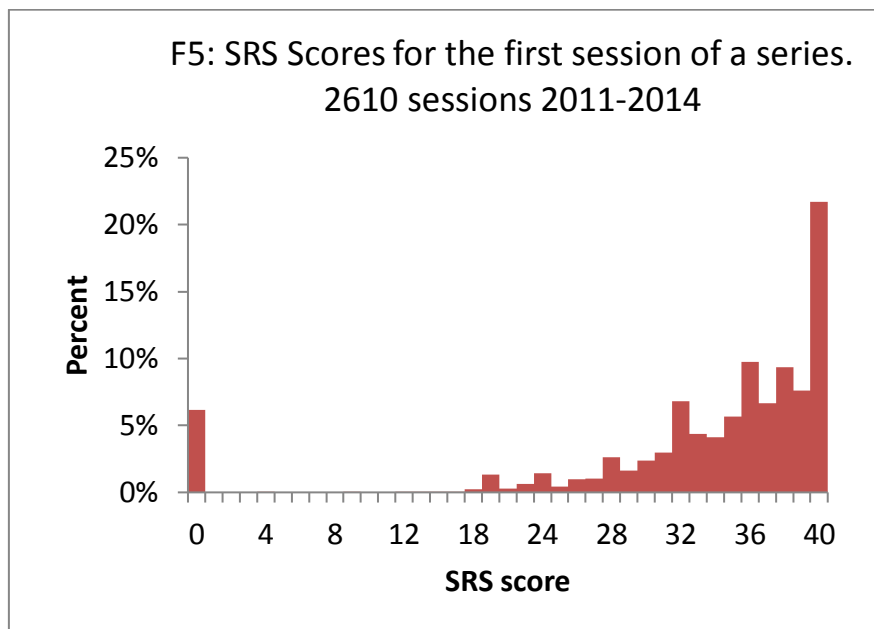
The same ORS pattern, in which scores that are a multiple of 4 are generally much more frequent than the surrounding data, is also seen in SRS scales in figure 3, most likely for the same reasons. Table 2 shows again how similar the means and standard deviations of the subscales are. An ANOVA confirms there are no significant differences between the means ($p=0.97$) and the intercorrelations at $r=0.77$ to 0.83 are very high and significant ($p=0.000$).

T2: SRS subscale scores for 1165 single session clients		
Subscales – “Your experience of – “	mean	SD
Relationship (with the counsellor)	8.01	2.83
Goals and Topics	7.98	2.82
Approach or Method	8.00	2.82
Overall	8.01	2.83

Figure 4 shows the distribution of ORS scores for clients who have attended more than one session in the first session. The distribution is very similar to the single session data in figure 2 and has the same mean (23) and very similar standard deviations (10.94 vs 10.50). The data has the same peaks at multiples of four as well. The percent scoring 0 is, however, about half that of the single session clients, suggesting possibly a greater level of preparedness to carry on at the outset.



When we look at the distribution of SRS scores for the first session in a series (figure 5) the data is very similar to the pattern for a single session (figure 3). There are 50% fewer 0 scores, and the mean at 32.6 is slightly higher (32.0 for the single session). This difference is significant (one tailed t-test for unequal variances; $p=0.049$) showing that group that goes beyond the first session was slightly more satisfied with it than the single session group.



The subscale data for both the ORS and SRS for those who continued with therapy has the same first session pattern as those who just had a single session – almost identical subscale scores that had no significant differences. This suggests that at least in the first stage of using the ORS and SRS that clients are unable to make subscale distinctions.

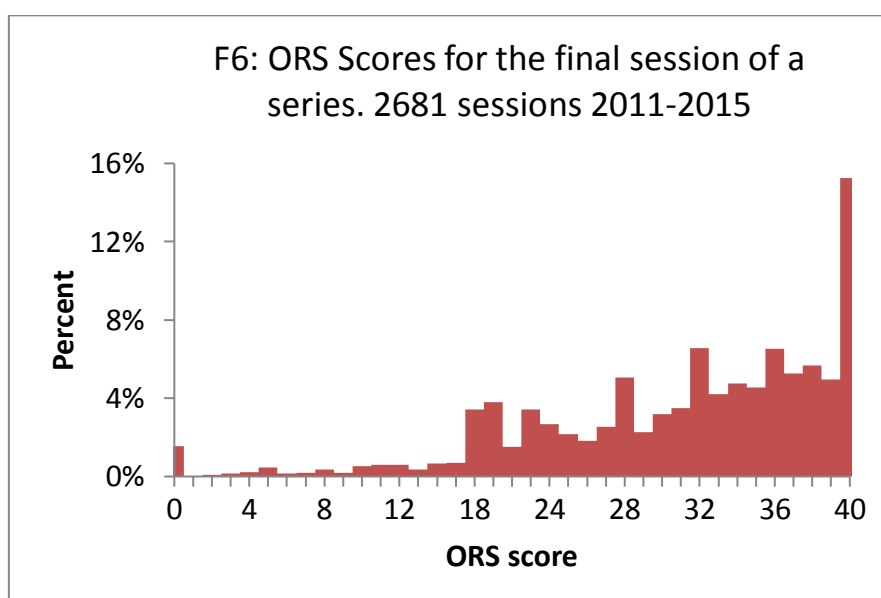


Figure 6 looks at the ORS scores for the final session counselling received by clients, all of whom began between January 1st 2011 and 31st December 2014, but about 10.5% continued therapy into 2015, with 2.3% having a counselling in April, the last available period in this dataset. Compared

with the first session data (figure 4), the distribution of final session scores has a mean of 30.3 (vs 23.1), is less variable ($SD=9.0$ vs 10.5) and the difference between the first and final sessions is highly significant (paired sample t-test, $p=0.000$). The median score was 33, 64% scored 30 or greater, the mode was 40 (15.3%) and only 1.6% scored 0.

Table 3 gives the subscale scores for the ORS and again the means and SDs are almost identical and intercorrelations very high and significant (all are $r=0.86$, $p=0.000$). There is no significant difference between the subscale scores (paired sample t-tests, $p=0.13$ to 0.88), and the approximately 1.8 point improvement in all subscale scores between the first (table 1) and final sessions is highly significant (paired sample t-tests, $p=0.000$) for all comparisons. This again asks the question whether the subscale scores are capable of differentiating the different components of mental well-being.

T3: ORS subscale scores for 2680 final session clients		
Subscales – “How are you feeling – “	mean	SD
Individually personal wellbeing	7.60	2.35
Interpersonally family close relationships	7.59	2.40
Socially Work school friendships	7.58	2.38
Overall	7.58	2.37

Comparing final session SRS scores (figure 7) with the first session (figure 5) we have mean of 35.2 (vs 32.6), less variable data ($SD=8.1$ vs 9.8) and a difference between the first and final sessions that is highly significant (paired sample t-tests, $p=0.000$). The median score was 38, 76.4% scored 35 or greater, the mode was 40 (35.2%) and 3.7% scored 0.

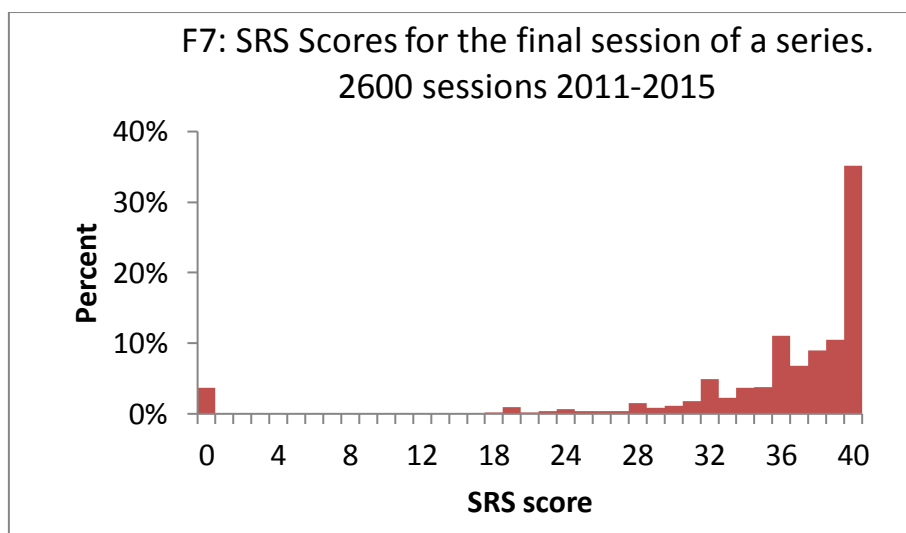
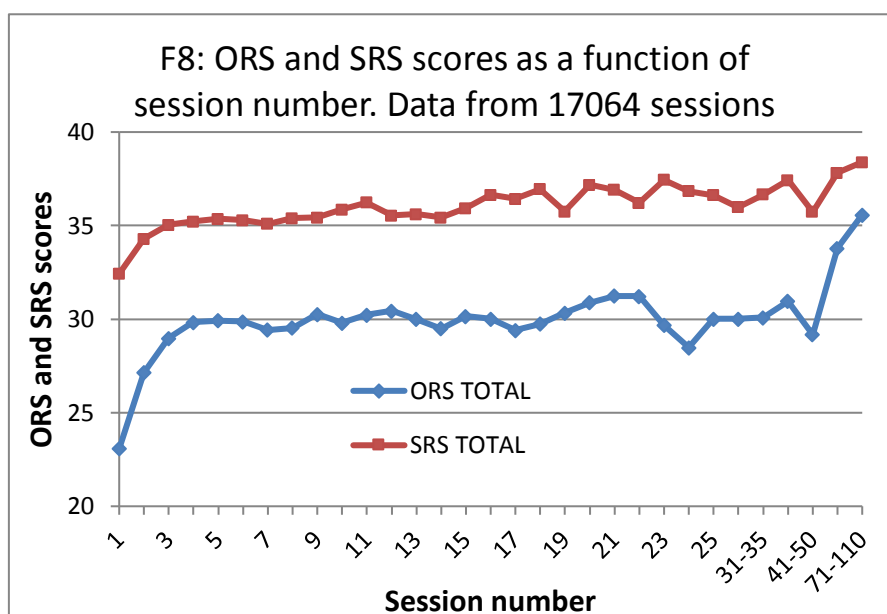


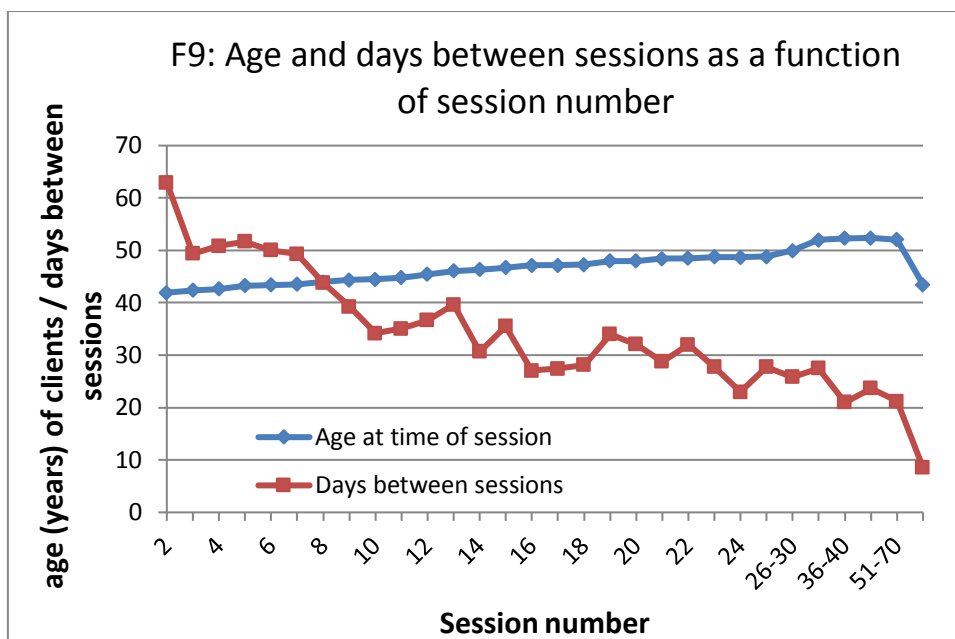
Table 4 shows the same pattern as table 3 of subscales scores having nearly identical means and standard deviations such that there is no significant difference between them (paired sample t-tests, $p=0.08$ to 0.91). As well, the approximately 0.79 improvement between first and final sessions is sufficient for there to be a highly significant difference (ANOVA, $p=0.000$) between all first and final session subscale scores. Again there a highly significant correlation ($r=0.43$, $p=0.000$) between ORS and SRS scores.

T4: SRS subscale scores for 2600 final session clients		
	mean	SD
Relationships	8.77	2.08
Goals and Topics	8.80	2.07
Approach or Method	8.79	2.05
Overall	8.81	2.14

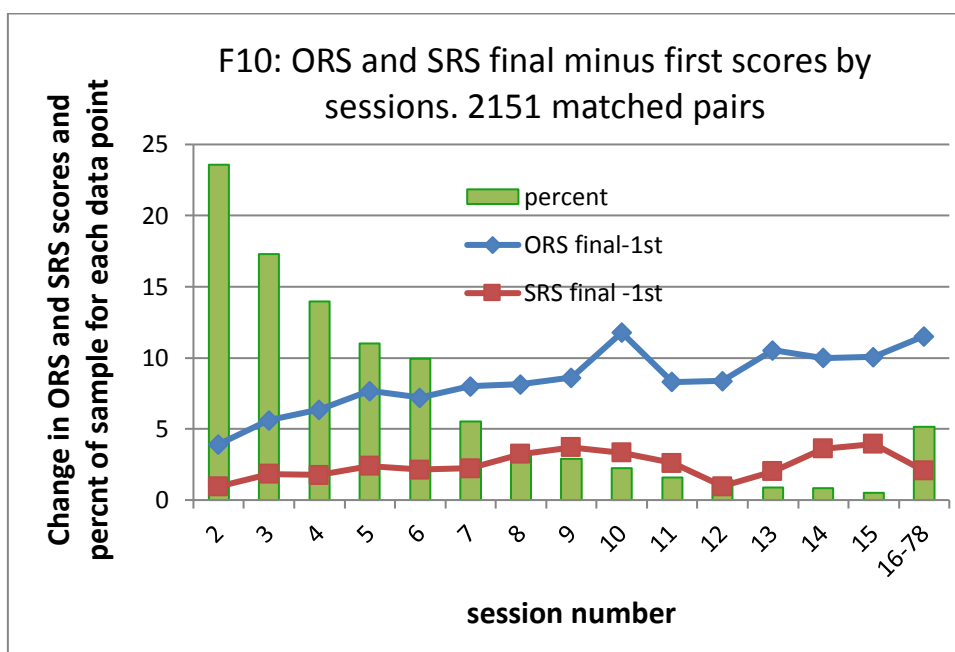
Next we consider in figure 8, the relationship between ORS and SRS scores and the number of counselling sessions. Both increase from starting points of 23.1 (ORS) and 32.4 (SRS) and then rise rapidly up to session 5 (29.9 and 35.4). The SRS hovers around 30 but spikes at 35.6 for nine very long-running participants (more than 51 sessions). This suggests a 7 point improvement in ORS as a result of counselling, but with little change after session 5. Overall, these changes are not significant (ANOVA, $p=0.078$). The SRS continues to climb another 3 points to 38.4, suggesting that the longer running clients are very positive about the counselling service (this is significant - ANOVA, $p=0.026$). Sessions 26-110 involved 151 clients – only 3.8% of the sample.



The average number of days between sessions is 47.3 or nearly seven weeks. Figure 9 shows the days between sessions as a function of session number. So the time between session 1 and session 2 is 60 days (8.5 weeks), and by session 10 this has fallen to five weeks, by session 16 to four weeks and by session 30 to around 3 weeks between sessions. The correlation days between sessions and number of session is significant ($r=0.094$, $p=0.000$). Figure 9 also looks at session number in relation to age and shows that the clients who are engaging in extended counselling (30 plus sessions) tend to be older by about 10 years that those doing one or two sessions (this significant, $r=0.18$, $p=0.000$). Of course doing 30 sessions will take about 30 months which will account for about 25% of the age difference.

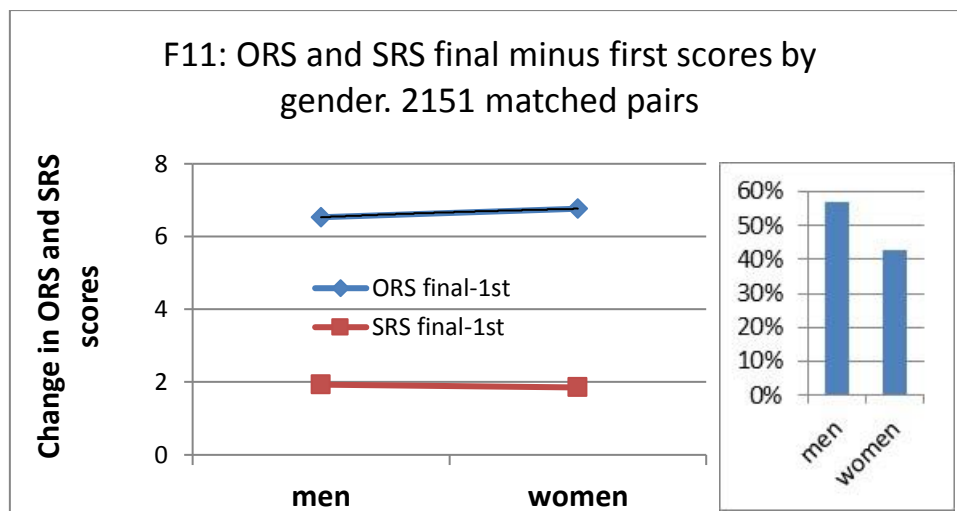


The finding above of ORS and SRS scores of 0 is problematic. For example, when we look at the 12 scores on the ORS for clients into their seventh or later session that are 0, we find that the SRS scores are also 0. Some of these clients are into their 15th, 16th, 18th, 44th or 110th session and it seems unlikely given their other data that 0 is a genuine reflection of their mental state or satisfaction with the counselling session. It is more likely that no data was taken and that instead leaving the data entry as blank, 0s were entered. It may be better to treat 0 totals for ORS and SRS as blank entries. In the next section of the analysis we remove all blank data and data fields where only 0s have been entered.



In figure 11 594 first and final sessions were removed because of an ORS or SRS total that was blank or 0. This left 2151 matched pairs where all data was present and first and final comparisons could be made. The same early 5 session spurt in the data occurs, showing 7.7 gain in ORS and a 2.4 gain in SRS, but both continue to gain with ORS peaking in session 10 with 11.8 gain and SRS peaking in

sessions 9 and 15 with gains of 3.7 and 3.9. The average gain in ORS is 6.6 and 1.9 for SRS. Both trends are highly significant ($p=0.000$ and $p=0.003$, ANOVA).



Men outnumber women by about 57% to 43%, but there is no significant difference in their ORS or SRS scores (ANOVA, $p=0.85$, $p=0.22$). The two transgender clients responded positively to counselling (2 and 3 sessions: ORS average change = 7.50; SRS average change = 8.00).

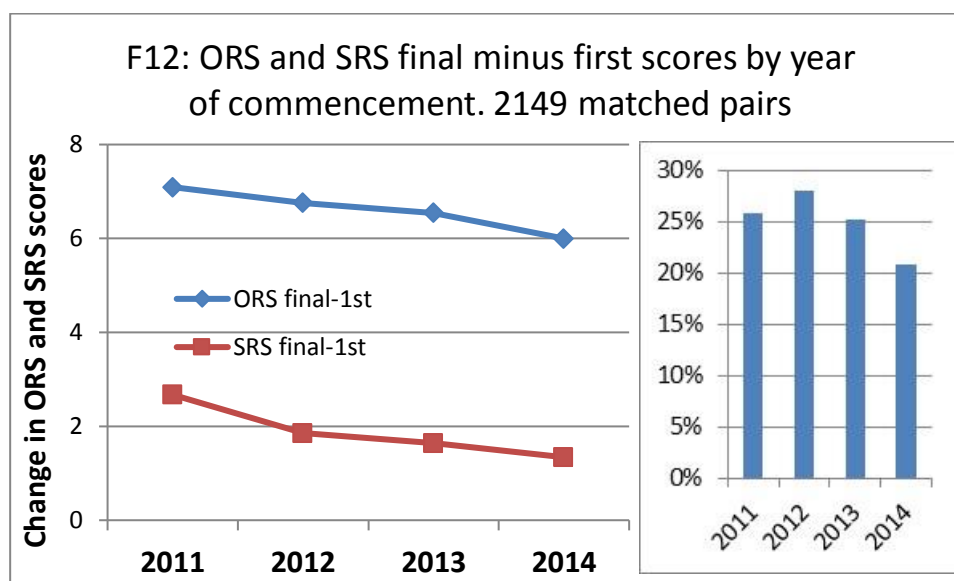
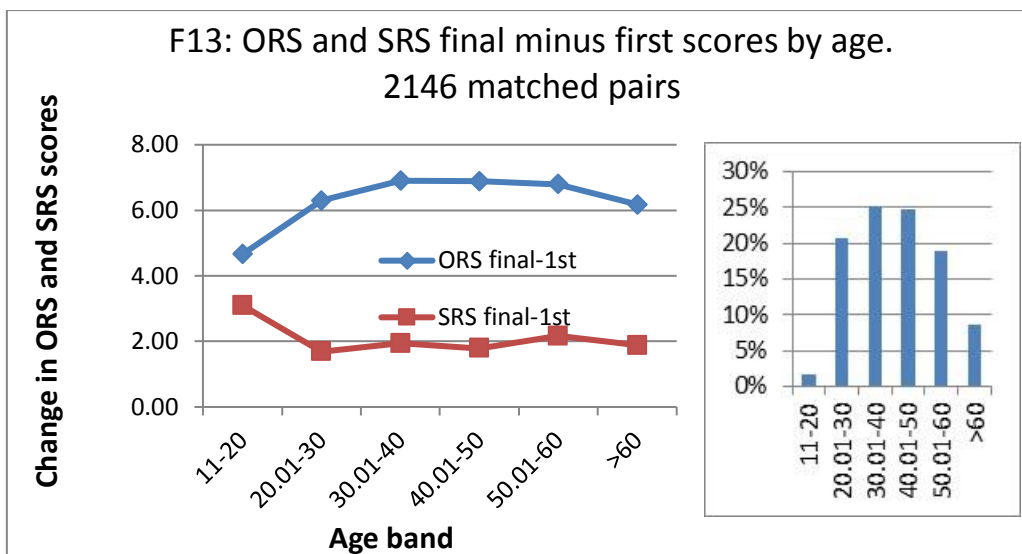


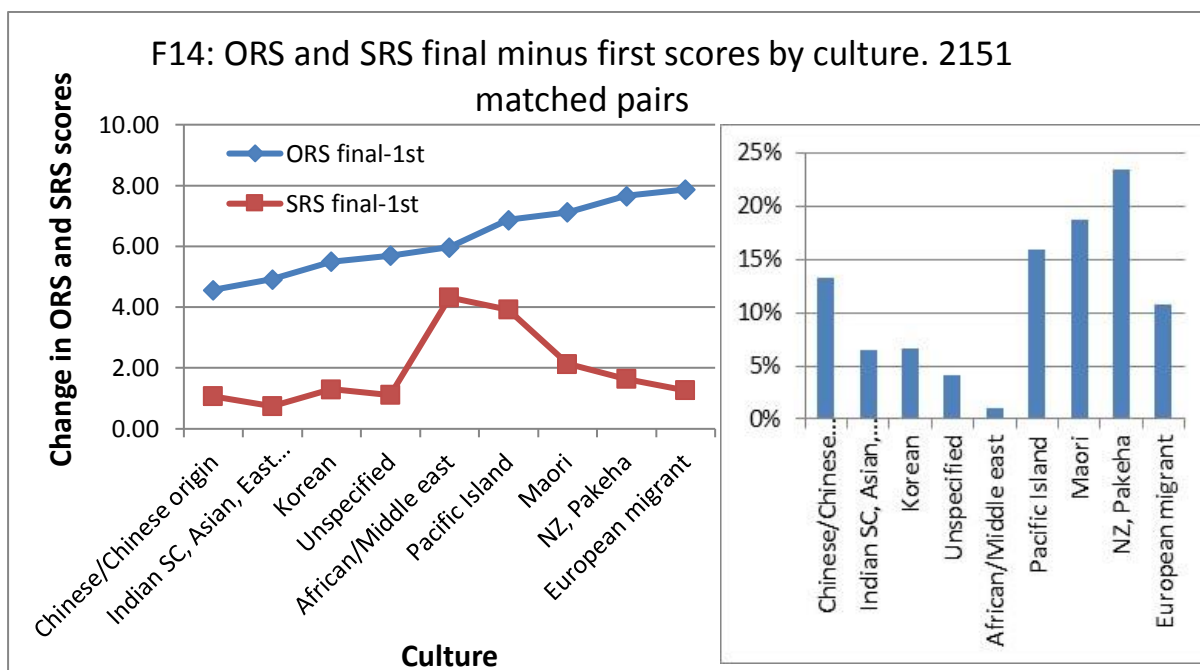
Figure 12 shows how the SRS and ORS scores have changed over the 4 years in the data sample. Firstly, the percent of this four year sample of PGF clients declined from 28% in 2012 to 20.1% in 2014. Over that time the ORS average change fell from 7.1 to 6.0 and the SRS average change from 2.7 to 1.3. The first vs final comparison is significant for all years (paired sample t-test, all $p=0.000$), and an analysis across 2011 to 2014 shows a significant fall in the SRS (ANOVA, $p=0.000$), but not for the ORS (ANOVA, $p=0.29$). While well-being improvement might not be falling across 2011 to 2014, satisfaction with counselling appears to be.

Figure 13 looks at ORS and SRS change (1st to final session) for six different age bands. Counselling appears to be most effective in the 30-50 age range where the average ORS change is 6.9, but falls away on either side with the youngest age group (only 1.7% of the sample) making the least

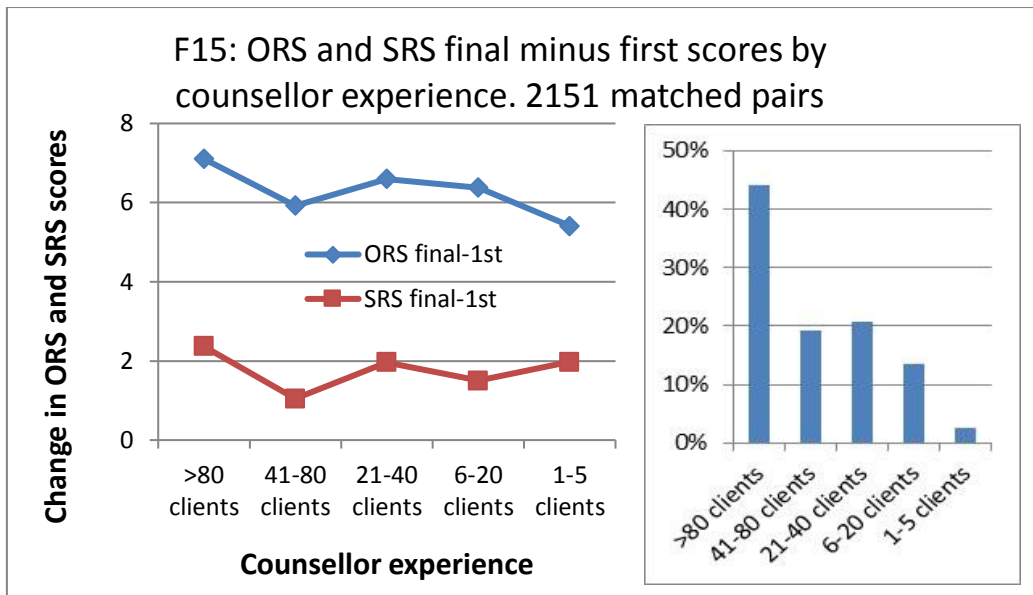
progress. However, this group also shows the greatest gains in satisfaction (SRS change). Neither the ORS nor SRS means are significantly different across age groups (ANOVA, $p=0.58$, $p=0.52$).



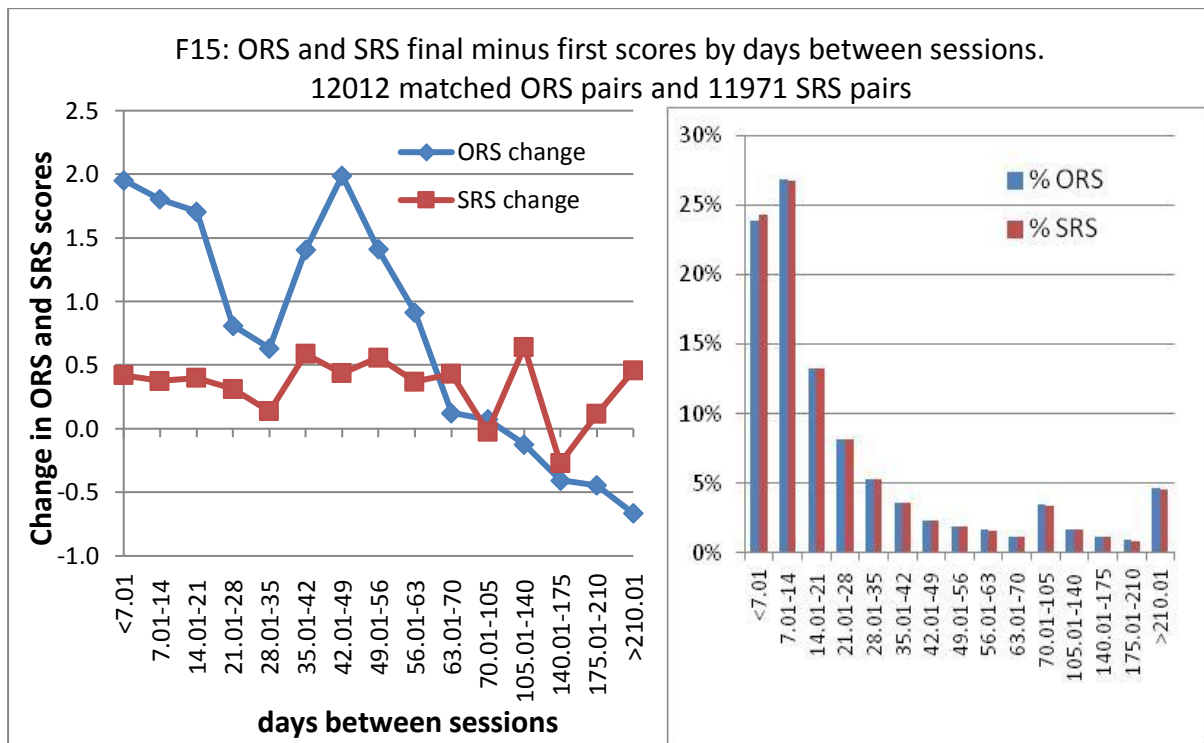
Culture appears to have a strong impact on service outcomes and satisfaction with services. At one end of the spectrum we have in figure 14, people of Chinese, Korean, Indian and Eastern Asian origin (about 25% of the PGF clientele) with an average ORS gain of 4.6 to 5.5, while at the other end there New Zealanders/Pākehā and migrant of European extraction (Including South America) scoring 2-3 points higher. These differences are highly significant (ANOVA, $p=0.000$).



Māori and Pacific Island clients, making up more than a third of PGF clients, also make strong gains on ORS scale, while also showing (along with a small group of African and Middle Eastern clients) the strongest positive shifts in satisfaction with services. SRS scores also show significant differences across cultures (ANOVA, $p=0.000$).



Very experienced counsellors (having more than 80 clients over the 4-year period of the analysis – 43% of all clients) do seem to make greater gains with clients (7.1 points on the ORS and 2.4 on the SRS) particularly when compared the small group of inexperienced counsellors (1-5 clients – 5.4 points gain on the ORS and 2.0 on the SRS). However, for the group of counsellors in the middle supporting between 6 and 80 clients each (53% of all clients), their level of experience is not a guide to progress made, and while there are significant differences in SRS change scores (ANOVA, $p=0.000$) there are not in ORS change scores ($p=0.19$)



As already stated the average break between sessions is nearly 7 weeks, and such a space between sessions means that therapy may lack continuity and therefore power. The figure above (figure 15)

explores the relationship between time between sessions and change on ORS and SRS scores between sessions, and does show a strong result for both ORS and SRS at 7 weeks, but this is against the general trend of the ORS data. Over the first weeks (covering 77% of all clients), ORS improvement drops by two thirds from 1.95 to 0.63. After the 7 week spike there is a further rapid decline, such that if there is ten weeks between sessions, no progress can be expected and for greater delays, people are losing ground. There is an $r=-0.079$ correlation between time between changes and change in ORS scores between session which is significant ($p=0.000$). This is not the case for SRS as $r=0.01$ and is not significant. As can be seen from figure 15, satisfaction with the session can continue to improve, even after long breaks.

This overview shows that substantial progress is made by the great majority of PGF clients and that they are highly satisfied with the services they receive. There are a number of issues that could usefully be addressed at this stage to further improve outcomes.

1. *A review of the way that SRS and ORS chart completion is done, scored and reviewed.* The lack of discrimination in the data between the ORS and SRS sub-scales suggests that clients pay little attention to the differences in emphasis between the subscales, which in turn suggests that the data from the ORS and SRS is not discussed with clients and they are not encouraged to make distinctions. Additionally, while much of the ORS and SRS scoring involves fine grain measurement (i.e. decimal points) much of it does not, and suggest that more care could be given to the recording of ORS and SRS scores. The number of blank entries and 0 scores requiring the discard of about 20% of the data reinforces this suggestion.
2. *A review of the difference outcomes for different cultural groups.* Table 5 below summarises the influence of the key demographic variables of age, gender and culture on ORS and SRS change scores. Only culture appears to be a predictor of outcomes and figure 14 shows ORS change scores to be 71% better for the best performing group than for the worst. If this result is validated in other ways, then it is an area where remedies should be sought.
3. *A review of the possible decline in service outcomes over the past four years.* Table 5 shows that there is a correlation between time of commencement of counselling with PGF and outcomes. ORS change scores have dropped by 18% and SRS by 50% from 2011 to 2014. Is this an artefact or a systemic issue.
4. *A review of structure of session delivery over time.* Its clear that the more sessions clients have access to (up to 10 sessions), the better the outcomes. Table 5 suggests that this is, by far, the strongest predictor of progress and satisfaction with services. Added to this is evidence that longer breaks between sessions lead to lower and in extreme cases negative ORS change scores. Is increasing the number of sessions and reducing the time between them a useful way to improve outcomes?
5. *A more detailed analysis of the role of counsellor experience.* There are a number of questions that could be asked here, such as the change of outcomes over time for experienced counsellors. A quick look at the data shows the ORS change scores drop by about a third for the most experienced counsellor group over 2011 to 2014.
6. *A review of the relationship between PGF ORS and SRS change scores and those found in the literature.* How good are these results compared similar services?

T5: Summary of the relationship between ORS and SRS change scores and key demographic and service delivery factors.			
Data from final-first session comparison. N=2152	ORS change	SRS change	
<i>Age session 1</i>	0.008	0.013	The age of clients on entry to and exit from the service has no significant correlation with change scores.
<i>Age final session</i>	0.012	0.022	
<i>Age group session 1(1=11-20, 2=20.001-30,5=50.001-60, 6=>60)</i>	0.010	0.010	
<i>Gender 0=male, 1=female</i>	0.012	-0.008	Gender has no correlation with ORS and SRS
<i>Chinese and Chinese origin</i>	-0.089**	-0.064**	Clients of Chinese. Indian sub-continent or South East Asian origin make less progress on ORS when compared to all other groups. New Zealanders, Pākehā and people of European origin make more progress. The satisfaction scores (SRS) of clients of Chinese. Indian sub-continent, South East Asian and European origin, show less improvement than for other groups
<i>Indian SC, South East Asia</i>	-0.050*	-0.060**	
<i>Korean</i>	-0.033	-0.032	
<i>Unspecified</i>	-0.022	-0.033	
<i>African, Middle Eastern</i>	-0.007	0.048	
<i>Pacific Islands</i>	0.011	0.173	
<i>Māori</i>	0.025	0.021	
<i>New Zealanders, Pākehā</i>	0.064**	-0.028	
<i>European origin</i>	0.046*	-0.042*	
<i>Counsellor experience – number of clients 2011-2014: 1=1-5, 2=6-20, 3=21-40, 4=41-80, 5=>80</i>	0.036	0.050*	
<i>Year 2011, 2012, 2013, 2014</i>	-0.041	-0.091**	Clients enrolled recently (up to the end of 2014) have less positive SRS and ORS change scores than those enrolled further back in time
<i>Date of session 1</i>	-0.042*	-0.090**	
<i>Series final position raw, 1-17</i>	0.179**	0.064**	Both ORS and SRS scores show greater improvement as a function of the number of counselling sessions a client receives. Number of sessions is the most powerful predictor of ORS improvement. The most powerful predictor of SRS improvement is the length of time (days of engagement) in counselling. However, average days between sessions is not predictor for either ORS or SRS.
<i>Series final position grouped after 10, 11-15=11; 16-20=12, etc</i>	0.219**	0.096**	
<i>Days of engagement</i>	0.059**	0.143**	
<i>Average days between sessions</i>	-0.029	0.111	
Data from all possible between sessions comparisons. N= 12012 matched ORS pairs and 11971 SRS pairs			
<i>Days between sessions</i>	-0.079**	0.009	Longer periods between sessions predict lower ORS change scores between sessions.

** . $r > 0.056$ or < -0.056 , $p < 0.01$ (2-tailed), * . $r > 0.41$ or < -0.41 . $p < 0.05$ (2-tailed).

Reese et al (2009) provide a summary, paraphrased below:

- No change (less than a gain of 5 points). For a client that has not shown reliable change (a gain of 5 points) after three sessions, practitioners are directed to address the alliance and the course of treatment or service provided. If the client has not demonstrated reliable improvement after six sessions, consultation with peers, and reviewing the service in supervision is suggested.
- Deteriorating (scores go down at least 5 points). Clients in this category are considered to be at-risk for ending the service prematurely or having a poor outcome. Practitioners should discuss possible reasons for the drop in score with the client, review the SRS items with the client to assess the alliance or consider changing the approach, frequency, mode, or even practitioner if no improvement is noted after three sessions.
- Reliable change (a gain of 5 or more points). The treatment or service is going accordingly. Practitioners are advised to reinforce changes and to continue treatment or service until progress begins to plateau, then a practitioner should consider reducing the frequency of sessions.
- Clinically significant change (at least a 5 point gain and crossing the ORS cut-off score of 25 during treatment). The client is probably no longer struggling with issues that led to him or her seeking services. Practitioners are advised to consolidate changes, anticipate potential setbacks, and to consider reducing the frequency of sessions.

Toward the end of every session, the SRS is administered to the client and again scored by the practitioner. If the total score is below 36 or one of the items is below 9, the practitioner intervenes and inquires about the reason for the lower scores.

Reese, R.J., Norsworthy, L.A., & Rowlands, S.R. (2009). Does a continuous feedback system improve psychotherapy outcome. *Psychotherapy: Theory, Research, Practice, Training*, 46, 418-431.

1. Valid initial Outcome Rating Scale (ORS); Duncan, 2014 ! 35+: Invalid initial score – why come to see you? Ø ORS average, 500,000+ administrations: 18-20 Ü Goal: Less than 1/3rd over the Clinical Cutoff (25, 28, 32)
 2. Reliable change index (RCI*) 6+ point ORS increase from the initial score
 3. Clinically significant change index (CSCI*) 6+ and cross the Clinical Cutoff Point
- Duncan (2014)

Schuman, D., Slone, N., Reese, R.J., & Duncan, B. (2014). Using client feedback to improve outcomes in group psychotherapy with soldiers referred for substance abuse treatment. *Psychotherapy Research*, <http://dx.doi.org/10.1080/10503307.2014.900875>

<https://heartandsoulofchange.com/> pre and post very similar to our ninth.