Do elite endurance athletes report their training accurately?

Øystein Sylta, Espen Tønnessen, and Stephen Seiler

This is the author’s version of an article published in the journal:

Do elite endurance athletes report their training accurately?

This is an original investigation with 4 tables and 3 figures. The abstract consists of 248 words and the text 3458 words.

Øystein Sylta¹, Espen Tønnessen² & Stephen Seiler¹

1) Faculty of Health and Sport Sciences, University of Agder, Kristiansand, Norway
2) The Norwegian Olympic Federation, Oslo, Norway

"This manuscript has been read and approved by all the listed co-authors and meets the requirements of co-authorship as specified in the Human Kinetics Authorship Guidelines. This manuscript is original and not previously published, nor is it being considered elsewhere until a decision is made as to its acceptability by the IJSPP Editorial Review Board."

Address for correspondence:

Øystein Sylta

Universitetet i Agder, Fakultet for Helse og Idrettsvitenskap,

Postboks 442, 4604 Kristiansand

Norway

E-mail: oystein.sylta@ui.no

Telephone 0047 92252792

Fax number 0047 38141001

Running head:

Validity in self-reported training
Abstract
Purpose: The purpose of this study was to validate the accuracy of self-reported (SR) training duration and intensity distribution in elite endurance athletes. Methods: Twenty-four elite XC skiers (25 ± 4 yr, 67.9 ± 9.88 kg, 75.9 ± 6.50 mL·min⁻¹·kg⁻¹) SR all training sessions during a ~14 d altitude training camp. Heart rate (HR) and some blood lactate (La) measurements were collected during 466 training sessions. SR training was compared to recorded training duration from HR monitors, and SR intensity distribution was compared with expert analysis (EA) of all session data. Results: SR was nearly perfectly correlated with recorded training duration (r = .99), but SR was 1.7 % lower than recorded training duration (P < .001). SR training duration was also nearly perfectly correlated (r = .95) with recorded training duration > 55 % HR_max, but SR was 11.4 % higher than recorded training duration > 55 % HR_max (P < .001) due to SR inclusion of time < 55 % HR_max. No significant differences were observed in intensity distribution in zones 1-2 between SR and EA comparisons, but small discrepancies were found in zones 3-4 (P < .001). Conclusions: This study provides evidence that elite endurance athletes report their training data accurately, although some small differences were observed due to lack of a SR “gold standard”. Daily SR is a valid method of quantifying training duration and intensity distribution in elite endurance athletes. However, additional common reporting guidelines would further enhance accuracy.

Key words: Validity, self-report, expert analysis, XC skiers, heart rate.
**Introduction**

Recently, the training characteristics of elite runners, rowers, cyclists and cross-country (XC) skiers have been described with a focus on basic aspects of training volume and intensity distribution over time-frames from weeks to an entire season.\(^1\)-\(^{12}\) The key method for quantifying training characteristics is self-reported (SR) training in diaries.\(^5\),\(^8\)-\(^{13}\) Such data may be used to examine the relationship between training dose and training adaptation alongside performance, and serves as a basis for mechanistic hypothesis generation. This, however, requires that SR diaries are valid with regard to activity form, volume, intensity distribution and frequency of training. In a validation study of SR training duration in recreational athletes, Borresen & Lambert\(^14\) conclude that quantification of an athlete’s actual training volume may be inaccurate when relying exclusively on SR data. However, this study was not conducted with elite athletes under rigorous conditions. Validation of individual session duration and total training volume is seemingly straightforward, but validation of intensity distribution is more challenging, both conceptually and operationally.

One approach is to continually register heart rate (HR) during each session, and use standardized or test profile based HR zone cut-offs to allocate heart rate *Time In Zone* (TIZ) to each intensity zone independent of power or pace.\(^6\),\(^7\),\(^13\),\(^15\) An alternative and commonly used method for SR intensity distribution among elite athletes is described as a “modified” *Session Goal* (SG) heart rate analysis in the literature\(^13\) and employed in several recent studies.\(^8\)-\(^{11}\) The SG heart rate method refines the TIZ method by using the primary goal of the session as a starting point for analyzing the intensity of the intended or core portion of each training session (steady state, threshold training or interval training). This method can be used as an alternative approach to TIZ heart rate analysis, or as in the original SG method\(^13\), a basis for a categorical allocation of each whole training session to an intensity zone, with or without corroborating perceived effort quantification.\(^13\),\(^16\) Validating intensity distribution gives rise to several challenges due to inconsistent methods and the absence of a common accepted “gold standard”.

Despite the importance of SR training data in describing endurance training best practice and developing testable training hypotheses, we have failed to identify previous studies validating the accuracy of SR training data provided by elite level athletes regarding session duration or intensity distribution. Therefore, the primary aim of the present study was to quantify the accuracy of SR training duration and intensity distribution among elite endurance athletes under rigorous altitude training-camp conditions.

**Methods**

**Subjects**

Twenty-nine elite XC skiers, aged 20-32 years, volunteered to participate in the study, which was approved by the Regional Ethics Committee of Southern Norway. All subjects provided informed written consent prior to participation. The 29 subjects were all athletes selected for the Norwegian National Team. Of the study participants, 10 athletes had won medals from senior World- or Olympic championships. Of the remaining 19 athletes, 18 had won medals from junior World championship or
placed among the top 3 in World Cup events. In the following XC skiing World Championship (Val di Fiemme, 2013) four months after this data collection period, the athletes included in this study between them won 7 gold, 4 silver and 5 bronze medals. Five of the subjects reported their intensity characteristics in a manner inconsistent with other 24 (details described below), and were excluded from the final analysis. The physical characteristics of the 24 subjects included in the present analyses are presented in Table 1.

(Intable 1)

Intensity zone determination

The Norwegian Olympic Federation (OLT) employs a five-zone aerobic intensity scale to prescribe and monitor training for endurance athletes. This scale is a general guideline, and the different zones are primarily based on lactate (La) and HR ranges (Table 2).

(Intable 2)

Intensity zone validation requires a reference standard for each athlete. In the present study, five aerobic intensity zones in line with OLT’s recommendations were determined prior to data collection for each athlete based on self-reported HR and La values defining individual intensity zone cut-offs. All athletes were well familiarized with the five-zone reference scale, having used this scale since junior age. Individual adjustments to HR and La values were performed based on experience and knowledge of each athlete’s own physiological characteristics. In addition, the SR intensity zones were verified against HR and La values acquired using standardized, on-site treadmill testing during the data collection period. Although HR monitors and La measurements have been found to provide accurate measures during physical activity, factors such as day-to-day variation, training status, training form, activity form, environmental conditions, diurnal variation, hydration status, altitude and medication may influence the relationship between work load and HR/La values. The athletes’ SR intensity-zones were therefore used as a reference standard, as opposed to laboratory testing results, which are performed under conditions not identical with training, and only have relevance for the zone 2-3 and zone 3-4 boundaries.

Registration of training

Validation was performed during an altitude training camp in Val Senales (Italy) October 2012, and average length of the data collection period was 14 days (range 8-18 days) (Table 3). During the period, six of the athletes contracted an illness or injury lasting two or more days. Athletes carried out their normal training and were instructed to follow their coaches’ recommendations. Training methods or organization was not discussed with the athletes during the data collection period. The
athletes were blinded to our aim to quantify SR training validity, and instead told that this was part of
data collection to monitor team training characteristics. Athletes were provided with detailed written
and verbal instructions via a group meeting explaining the importance of keeping their training
diaries and using a HR monitor during every training session.

**Self-report**

Due to concerns about internet access stability, athletes were provided with simple hard copies of
their normal on-line training diary, and asked to record their daily training information after each
session as per their normal routine. The information in the diary consisted of quantifiable data
regarding activity form, duration, intensity distribution and comments. The majority of athletes (24
of 29) divided the total duration of each session into intensity zones, based on a modified SG
approach where objective information from their HR watches, La` measurements and stress
responses were used to determine relevant zones. Five athletes transferred their HR watch data
directly into software and recorded TIZ analysis from software analysis as SR. These five athletes
were excluded from all analyses for consistency. There were also some differences in registration
protocols for interval sessions. Some athletes included the recovery time between interval work
bouts as moderate or high intensity (zone 3-5), whereas others logged it as training time in zone 1.

**Recorded training duration**

All athletes used Garmin HR watch Forerunner 910XT or 610 (Garmin, Kansas, US) for every session.
HR sampling frequency was one hertz. HR data were uploaded to Garmin Training Center (ver. 3.6.5)
and further analyzed in Microsoft Excel (2010). In total, 466 of 530 sessions (88 %) were analyzed
with recorded HR data. Data from the remaining 12 % of sessions was excluded due to incomplete
HR data.

SR training duration was compared with “actual” training duration from complete HR records via two
methods. First, we compared SR session duration with the total recorded session duration retrieved
from HR files. Secondly, we restricted “actual” training duration to include only HR values > 55 % of
HRmax (typically HR > ~ 100). The rationale for this second analysis was that training with lower HR
than 55 % HRmax is below the OLT recommendation for zone 1 (Table 2).

**Expert analysis intensity distribution**

Validation of SR intensity distribution was achieved by comparison of SR data from athletes with
individual analysis by investigators of all available data for each training session. This analysis was
termed expert analysis (EA). The EA method is based on the previously described modified SG
analysis, combined with HR and La` measurements. Sessions performed in zones 1 and 2 were
defined using HR curves as a starting point, and then categorized into time in different zones in an
appropriate manner. Sessions in zones 3, 4 and 5 used the primary goal of the session’s core section,
alongside HR and La\textsuperscript{-} values to distribute the training time into the appropriate intensity zones. Recovery phases in interval sessions (zones 3-5) were categorized as zone 1 or 2, depending on the external load during that phase. EA included allocation of HR values < 55 \% HR\textsubscript{max} to match total SR time, categorized as zone 1.

**Statistical analyses**

Total training volume was calculated as the total duration of endurance, strength, sprint and plyometric training. Endurance training was further categorized into five intensity zones. In analyses of training SR validity, only the endurance portion of total training time was included.

Training characteristics data are reported as mean ± standard deviation (SD) and range. Pearson's Product Moment correlation was used to quantify the relationship between SR and HR based recorded training duration. Correlation magnitude \((r)\) was interpreted categorically as small \((r\) from 0.1 to 0.3), moderate \((r\) from 0.3 to 0.5), large \((r\) from 0.5 to 0.7), very large \((r\) from 0.7 to 0.9) and nearly perfect \((r\) from 0.9 to 1.0) using the scale presented by Hopkins et al.\textsuperscript{19} A paired sample T Test was used to identify systematic differences between the methods, and the 95 \% confidence intervals (CI) bounding the difference were calculated. The limits of agreement between SR and recorded training duration were calculated using a Bland-Altman plot.\textsuperscript{20}

All statistical analyses were performed using SPSS 18.0 (SPSS Inc, Chicago, IL, USA) and MedCalc (ver. 12.4.0.0), and statistical significance was accepted at the \(P < .001\) level.

**Results**

**General training characteristics**

General training characteristics based on 466 SR training sessions during the altitude camp are shown in Table 3. All sessions were either endurance sessions, or endurance sessions including strength, sprints or plyometric. Each training day typically consisted of two sessions. AM sessions were primarily on-snow skiing on a glacier ~3000m above sea level, and PM sessions were primarily roller-skiing or running in the valley near Val Senales (1200-2200m above sea level).

(Table 1)

**Accuracy in self-reported training duration**

There was a nearly perfect correlation \((r = .99; P < .001)\) between SR and HR watch registered training duration in each session \((n = 466)\) (Figure 1, PANEL A). SR training duration \((117 ± 36\text{ min})\)
was slightly but significantly lower than training duration derived from HR recordings (119 ± 37 min),
(98.3 ± 6.4 %; 95 % CI: 97.7 - 98.9; P < .001).

(Figure 1)

Figure 2/PANEL A shows the Bland-Altman plot of SR and recorded training duration in each session
(n = 466). The limits of agreement were -2.7 to -1.7 min. The variation around the mean difference
(-2.2 min) appeared to be random. Among all sessions, 77 % were within +/- 5 min deviation between
SR and recorded values.

(Figure 2)

There was a nearly perfect correlation (r = .95; P < .001) between SR and HR watch registered
training duration > 55 % HR_{max} in each session (n = 466) (Figure 1/PANEL B). However, under training
camp conditions, athletes systematically “over reported” the duration of training time that their HR
exceeded 55 % HR_{max}. Averaged SR training duration (117 ± 36 min) was significantly higher than
training duration derived from HR recordings > 55 % HR_{max} in each session (106 ± 34 min), a
difference of 11.4 ± 13.5 %, (95 % CI: 10.3 – 12.5 %; P < .001). The mean difference in SR versus
recorded training duration > 55 % HR_{max} was 10.7 min; CI: 9.7 – 11.8; P < .001 (Figure 2/PANEL B).

Intensity distribution

SR training volume was not significantly different from EA allocations for intensity zones 1 and 2
(Table 4 & Figure 3). Compared with EA based distributions, the athletes’ SR method significantly
overestimated total time spent in zone 3 during the training camp by 37 ± 25 min (1.7 ± 0.9 %) (P < .001),
while underestimating total time spent in zone 4 by 11 ± 12 min (0.4 ± 0.4 %) (P < .001). During
the entire camp, no training time in zone 5 was detected via EA and none was identified by SR.

(Figure 3)

Discussion

The main finding of this study is that elite endurance athletes accurately SR their training data. SR
training duration closely matches verified duration derived from HR recordings. Furthermore, the SR
intensity distribution is also accurately reported, although there are slight differences between zones
compared to EA.
We chose to perform data collection under very rigorous conditions during a high-altitude training camp. For Norwegian XC skiers, altitude training forms an important and routine component of the annual training regime with 60-100 days typically spent at altitude in the period September to February. We were also interested in examining the athletes’ intensity control during high intensity training sessions because training “incorrectly” at altitude can increase the risk of overreaching. However, it is worth noting that SR of training would likely have been even more accurate under normal sea-level conditions, particularly with regard to intensity distribution. Physiological parameters such as HR and La can respond somewhat differently in altitude, which in turn may influence the athlete’s perceived exertion and intensity control.

**Self-reported training duration**

During a ~14 day period, agreement between SR and HR recorded training duration was very high, with SR training duration in each session being 98.3 % of the training duration derived from HR recordings including HR values < 55 % HRmax. Overall, a nearly perfect correlation ($r = .99$) was found between SR and recorded duration in each session. No previously published studies have reported similar comparisons for elite athletes. Contradictory to our findings, Borresen & Lambert showed that recreational athletes’ quantification of training volume can be inaccurate when based on SR data alone. However, comparing the present study results directly with the results of Borresen & Lambert is unsuitable due to different methods.

In the present study athletes were instructed to use a HR watch during all training sessions and report their training daily. As such, it is reasonable to expect high accuracy in SR training duration data. Even a discrepancy of only 1.7 % may be viewed as noteworthy, and this difference was in fact statistically significant ($P < .001$). In practice, the discrepancy was due to athletes deducting a small percentage of time spent during each session to compensate for time that in reality cannot be counted as effective training time (drinking, urinating, very low intensity etc.). This would also explain the variation around the mean shown in Figure 2/PANEL A. However, this variation is small, with 77 % of all 466 sessions within +/- 5 min of the mean difference.

There was also a nearly perfect correlation ($r = .95$) between SR and recorded training duration > 55 % HRmax. However, under training camp conditions, athletes systematically over-reported training duration > 55 % HRmax by about 11 %. This indicates that a meaningful portion of reported training was performed at intensity below OLT’s recommended lowest limit for “effective” endurance training. If this practice were followed over the 800+ hours of these athletes’ typical annual training volume, the difference would extrapolate to ~100 hours of total training. However, the altitude “training camp” environment on the glacier probably exaggerates this overestimation of “effective” training time.
The over-reporting of training duration during a training camp can be partly explained by the norms and culture for recording training time in this specific XC skiing environment, where athletes keep their watches running during the entire session, even when stopping briefly for various reasons (e.g. hydration, urination and confer with coach). Other possible explanations are that HR may drop below 55 % HRmax when skiing downhill, or simply that the athletes (as instructed by their coach) deliberately train extremely slowly during initial training sessions at this altitude. In addition, the environment that athletes are exposed to during an altitude training camp may be viewed as atypical; coaches continuously providing feedback, physiologists measuring lactate and giving feedback on intensity, technique training sequences which include recovery phases, testing of a large number of skis, highly disciplined routines with regard to drinking every 20 minutes etc. To our knowledge, no previous studies to date have shown similar results. However, more studies, and during normal conditions, are necessary to corroborate these findings.

**SR intensity distribution**

There were almost no differences with regard to SR intensity distribution for zones 1 and 2, and only small differences for zones 3 and 4 compared to the EA. Zone 5 training was either reported in SR or detected by EA during any of the 466 sessions analyzed.

In the low intensity range (zones 1 & 2) no significant allocation differences were found, although athletes tended to SR some zone 2 training time to zone 1. Other than individual HR cut-offs for zones (Table 2), there are no clear physiological distinctions between zone 1 and 2, and in practical terms it may therefore be difficult for athletes to allocate total time in easy sessions between these two zones. Some athletes failed to record any time in zone 2, while others used HR data from watches to distribute total time. During EA investigators used HR-curves to allocate phases during a training bout where HR was clearly in zone 2. For these reasons we found some small, but not significant differences between SR and EA methods for zone 1 & 2.

In the high intensity range (zone 3-5), small but significant differences were found between SR and EA in zones 3 and 4. SR intensity distribution overestimated time in zone 3 and underestimated time in zone 4. During interval sessions most athletes registered recovery phases as time in the same high intensity zone as the effort (zone 3-4), while we as investigators did not, due to lower external load during that phase. This primarily explains why there is a difference in zone 3 in which the majority of interval sessions were conducted. In addition, we found a small discrepancy in zone 4. This was due to athletes not registering time in zone 4 despite HR and La values being in this zone for some intervals, particularly towards the latter part of sessions.

A limitation of this validation study was that no zone 5 training was prescribed during the altitude camp such that the full range of intensity distribution was not used during the training period.
However, SR of no zone 5 training was confirmed by EA throughout all 466 evaluated sessions, giving support to the validity of SR.

**Practical applications**

There is no SR “gold standard” and although we found some minor discrepancies between SR and recorded duration or EA intensity distribution, we suggest that these small differences are due to the absence of clearly-defined guidelines. Our findings indicate that scientists and coaches can rely on the validity of SR training data from elite endurance athletes, but common guidelines would further ensure the accuracy and comparability of SR data across individuals and sport disciplines.

For continuous training in zone 1 and 2, we recommend the use of HR values and external load to allocate periods clearly in different zones for SR in diaries. Furthermore, we suggest stopping watches in the case of obvious pauses during training (e.g. drinking, urinating etc.). For higher intensity continuous or interval sessions at or above the lactate threshold (zones 3-5 in the present study) we recommend distributing training time using a modified session goal approach, where the intended core portion of each session is used as the starting point for allocating zones, in combination with HR and La` values. While there are defensible arguments in both directions, we recommended that recovery phases during interval sessions are registered in zones corresponding to the actual external load. That is, an interval session of 5 x 8 min in zone 4 with 2 min recoveries would be recorded as 40 min of zone 4 training time, not 48 min. This will promote both internal consistency across zones, but also consistency throughout the season. Including recovery time between intervals as time in the high intensity zone can create a “false increase” in high intensity training if the rest to recovery ratio is changed as part of the periodization process.

To our knowledge, this is the first validation study investigating the SR of training information by elite level athletes. Additional work is needed in this area under routine training conditions and with different sports as a quality assurance platform for further research on optimization of the training process.

**Conclusions**

This study provides evidence that, overall, elite endurance athletes accurately self-report their training duration and intensity distribution. Common guidelines and a specific “gold standard” for training self-report may further increase validity.
References


**Table captions**

**Table 1**: Physical characteristics of the subjects, male: n=12, female: n=12.

**Table 2**: The Norwegian Olympic Federations five-zone intensity scale.

**Table 3**: General training characteristics based on SR training, n=24.

**Table 4**: Mean training time distribution within different zones based on SR and EA, n=24.

**Figure captions**

**Figure 1**: Relationship between SR and recorded training duration in each session (n=466). Dotted line indicates line of identity. PANEL A shows recorded training duration including HR values < 55 % HRmax, and PANEL B shows recorded training duration excluding HR values < 55 % HRmax.

**Figure 2**: Bland-Altman plot of SR and recorded training duration (n=466). PANEL A shows recorded training duration including HR values < 55 % HRmax, and PANEL B shows recorded training duration excluding HR values < 55 % HRmax.

**Figure 3**: Mean ± SD percentage of time spent in each of the five intensity zones (n=24). Open bars denote SR, while filled bars represent EA. EA - zone 1 includes HR values < 55 % HRmax matching SR. Panel A: zone 1, panel B: zones 2-5. *Paired Samples T Test, P ≤ .001.
<table>
<thead>
<tr>
<th></th>
<th>Male Mean ± SD</th>
<th>Female Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>24.8 ± 3.23</td>
<td>24.2 ± 4.19</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>179.8 ± 5.77</td>
<td>167.8 ± 5.36</td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td>75.6 ± 6.33</td>
<td>60.2 ± 5.97</td>
</tr>
<tr>
<td>HR\textsubscript{max} (beats \cdot min\textsuperscript{-1})</td>
<td>193.8 ± 8.33</td>
<td>194.8 ± 7.51</td>
</tr>
<tr>
<td>VO\textsubscript{2max} (mL \cdot kg \cdot min\textsuperscript{-1})</td>
<td>80.9 ± 3.70</td>
<td>70.9 ± 4.58</td>
</tr>
<tr>
<td>VO\textsubscript{2max} (L \cdot min\textsuperscript{-1})</td>
<td>6.1 ± 0.46</td>
<td>4.3 ± 0.32</td>
</tr>
<tr>
<td>Intensity zone</td>
<td>Lactate(^#) (mmol \cdot L(^{-1}))</td>
<td>Heart rate (% max)</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>5</td>
<td>6.0-10.0</td>
<td>92-97</td>
</tr>
<tr>
<td>4</td>
<td>4.0-6.0</td>
<td>87-92</td>
</tr>
<tr>
<td>3</td>
<td>2.5-4.0</td>
<td>82-87</td>
</tr>
<tr>
<td>2</td>
<td>1.5-2.5</td>
<td>72-82</td>
</tr>
<tr>
<td>1</td>
<td>0.8-1.5</td>
<td>55-72</td>
</tr>
</tbody>
</table>

Note. The reference values in this scale are guidelines only, and individual adjustments are required.
\(^\#\) Measured with lactate pro LT-1710
### Table 3

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD</th>
<th>Range (min – max)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recorded training days per athlete</td>
<td>13.3 ± 1.83</td>
<td>8 – 18</td>
</tr>
<tr>
<td>SR training volume (hours)</td>
<td>39.5 ± 9.40</td>
<td>25 – 60</td>
</tr>
<tr>
<td>SR endurance training volume (hours)</td>
<td>37.7 ± 8.86</td>
<td>24 – 57</td>
</tr>
<tr>
<td>Endurance training (%)</td>
<td>95.7 ± 1.83</td>
<td>92 – 100</td>
</tr>
<tr>
<td>Strength training (%)</td>
<td>3.4 ± 1.67</td>
<td>0 – 7</td>
</tr>
<tr>
<td>Sprint training (%)</td>
<td>0.8 ± 0.74</td>
<td>0 – 30</td>
</tr>
<tr>
<td>Plyometric (%)</td>
<td>0.1 ± 0.20</td>
<td>0 – 1</td>
</tr>
</tbody>
</table>
### Table 4

<table>
<thead>
<tr>
<th>Zone 1 (min)</th>
<th>Zone 2 (min)</th>
<th>Zone 3 (min)</th>
<th>Zone 4 (min)</th>
<th>Zone 5 (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SR</strong></td>
<td><strong>EA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2026 ± 497</td>
<td>1990 ± 461</td>
<td>109 ± 133</td>
<td>123 ± 51</td>
<td>7 ± 12</td>
</tr>
</tbody>
</table>

EA (zone 1) includes HR values < 55 % HR$_{\text{max}}$ matching SR. *Paired Samples T Test, $P \leq .001$
Figure 1

Panel A

Panel B
Figure 2

**PANEL A**

**PANEL B**
Figure 3

PANEL A

PANEL B