

Analysis of the performance tests (Objective Structured Clinical Examination) by using Item Response Theory (IRT) and Neural Test Theory (NTT)

**M Miyamoto, Y Mori, A Miyazaki, T Kubota,
H Yoneda, K Suzuki
Education Center, Osaka Medical College**

The Generalizability Theory has showed that the variance of correlations between examinees and stations is so larger than the inter-rator variance.

This means that scenario or station specificity is very important.

However, it will be very useful if we can get the common parameters among them.

We can compare the qualities of examinees and scenarios over years and places.

The **Graduation OSCEs** of Osaka Medical College for 6th year students have been performed from 2006 through 2009.

In the case of abdomen station, we have two Scenarios of “**abdominal pain**” and “**abdominal distension**” .



**The number of examinees for “abdominal pain” was a total of 272
68 in 2006, 72 in 2007, 68 in 2008 and 64 in 2009.**

**The number of examinees for “abdominal distension” was a total of 256
62 in 2006, 70 in 2007 and 64 in 2008 and 60 in 2009 results.**

We have used the same scenarios over the past 4 years.

Each scenario for “abdominal pain” and “abdominal distension” contained

45 items including

20 for medical interview (SHARED)

**25 for physical examination
(PARTLY SHARED)**

Each student's performance was judged by two assessors. A student can get "1" for each item just in the case that both raters admit.



| Table 1 Assessment items from interview for "abdominal pain" and "abdominal distension" | | "abdominal pain" | | "abdominal distension" | |
|--|--|----------------------------------|-------|----------------------------------|-------|
| Interview process | | average | SD | average | SD |
| 1) | Greeting | 0.981 | 0.138 | 0.934 | 0.249 |
| 2) | Selfintroduction | 0.986 | 0.120 | 0.990 | 0.101 |
| 3) | Confirmation of name | 0.976 | 0.154 | 0.974 | 0.158 |
| 4) | Listening to pt's complaints at opening | 0.966 | 0.181 | 0.974 | 0.158 |
| 5) | Appropriate, polite wording | 0.971 | 0.168 | 0.944 | 0.231 |
| 6) | Appropriate eye contact | 0.947 | 0.224 | 0.959 | 0.198 |
| 7) | Appropriate posture and attitude | 0.971 | 0.168 | 0.969 | 0.173 |
| 8) | Sympathetic words and attitude | 0.760 | 0.428 | 0.878 | 0.329 |
| 9) | Appropriate summary and confirmation | 0.933 | 0.251 | 0.918 | 0.275 |
| 10) | | | | | 5 |
| Passing Rate of Assessment items from interview for "abdominal pain" and "abdominal distension" | | | | | |
| 11) | | | | | 36 |
| 12) | Ask for the properties of the symptom | 0.957 | 0.204 | 0.990 | 0.101 |
| 13) | Ask for continuations of the symptom | 0.841 | 0.366 | 0.903 | 0.297 |
| 14) | Ask for the level of the symptom | Mean 0.900 (n=25) | | Mean 0.898 (n=25) | |
| 15) | Ask for the progress of the symptom | | | | |
| 16) | Ask for situations when the symptom occurs | | | | |
| 17) | Ask for factors in which it betters or worsens | | | | |
| 18) | Ask for accompanying symptoms | | | | |
| 19) | Ask the patient's response to the symptom | | | | |
| 20) | Ask the condition, and symptoms of the whole body. | 0.889 | 0.314 | 0.872 | 0.334 |
| 21) | Ask for the past history | 0.798 | 0.402 | 0.837 | 0.371 |
| 22) | Ask for the family history | 0.952 | 0.214 | 0.974 | 0.158 |
| 23) | Ask for extra information | 0.904 | 0.296 | 0.786 | 0.411 |
| 24) | Ask for the social situation | 0.851 | 0.357 | 0.806 | 0.396 |
| 25) | Ask for the pt's interpretation model | 0.875 | 0.332 | 0.745 | 0.437 |
| | Rating scale | 4.740 | 0.652 | 4.816 | 0.904 |

| Table 2 Assessment items for physical examination of "abdominal pain" and "abdominal distension" | | | | | | | |
|--|---|---------|-------|----------------------------------|----------------------|---------|-------|
| Scenario1 "abdominal pain" | | | | Scenario2 "abdominal distension" | | | |
| Consideration to patient | | average | SD | | | average | SD |
| 1) | Ask for physical exam and get agreement | 0.913 | 0.282 | 1) | Same as scenario1-1) | 0.918 | 0.275 |
| 2) | Imposing voice in each examination | 1.000 | 0.000 | 2) | Same as scenario1-2) | 0.985 | 0.123 |
| 3) | Comprehensible imposing of voice | 0.981 | 0.138 | 3) | Same as scenario1-3) | 0.985 | 0.123 |
| 4) | Consideration to pain | 0.990 | 0.098 | 4) | Same as scenario1-4) | 0.980 | 0.142 |
| 5) | Warming of the stethoscope | 0.942 | 0.234 | 5) | Same as scenario1-5) | 0.964 | 0.186 |
| Examination techniques | | | | | | | |
| 6) | Was the abdomen exposed enough? | 0.986 | 0.120 | 6) | Same as scenario1-6) | 0.918 | 0.275 |
| 7) | Was sexuality considered? | 0.976 | 0.154 | 7) | Same as scenario1-7) | 0.969 | 0.173 |
| 8) | Examine in order of inspection → auscultating → percussion → palpation? | 0.938 | 0.243 | 8) | Same as scenario1-8) | 0.898 | 0.303 |
| <Inspection> | | | | <Inspection> | | | |
| 9) | Inspection of important symptoms and expression of the findings | 0.928 | 0.259 | 9) | Same as scenario1-9) | 0.811 | 0.392 |

Passing Rate of Assessment items for physical examination of "abdominal pain" and "abdominal distension"

| | | | | | | | |
|---------------|---|-------|-------|---------------|---|-------|-------|
| 13) | confirmation of the painful area | 0.861 | 0.347 | 13) | Same as scenario1-13) | 0.868 | 0.316 |
| | | | | | Was the peritoneal fluid confirmed by accurate technique and the finding expressed? | 0.488 | 0.488 |
| | | | | | Were the upper and lower borders of dullness confirmed by accurate percussion? | | |
| | | | | | Was spleen enlargement confirmed by accurate percussion? | | |
| <Palpitation> | | | | <Palpitation> | | | |
| 14) | Were the lower limbs bent | | | 14) | Same as scenario1-14) | | |
| 15) | Proper superficial palpation on the whole abdomen first | | | 15) | Same as scenario1-15) | | |
| 16) | Was the deep palpation performed correctly next? | 0.938 | 0.243 | 16) | Same as scenario1-16) | 0.939 | 0.240 |
| 17) | Was the painful part (in the right low quadrant) palpated at the end? | 0.803 | 0.399 | 17) | Was the liver confirmed by accurate palpation? | 0.776 | 0.418 |
| 18) | Was the pain pressure point confirmed? | 0.745 | 0.437 | | | | |
| 19) | Was rebound tenderness confirmed? | 0.750 | 0.434 | | | | |
| 20) | Was muscular guarding confirmed? | 0.620 | 0.487 | | | | |
| Rating scale | | 4.808 | 0.793 | Rating scale | | 4.699 | 0.953 |

**Mean
0.891
(n=20)**

**Mean
0.877
(n=20)**

Uni-dimensionality

The **Item Response Theory** (IRT) needs **uni-dimensionality** strictly for the analyzed items.

NTT is a nonparametric test theory that uses a mechanism based on the self-organizing map (SOM).

However it seems better even for the Neural Test Theory (NTT) to take care of such a **uni-dimensionality** among items.

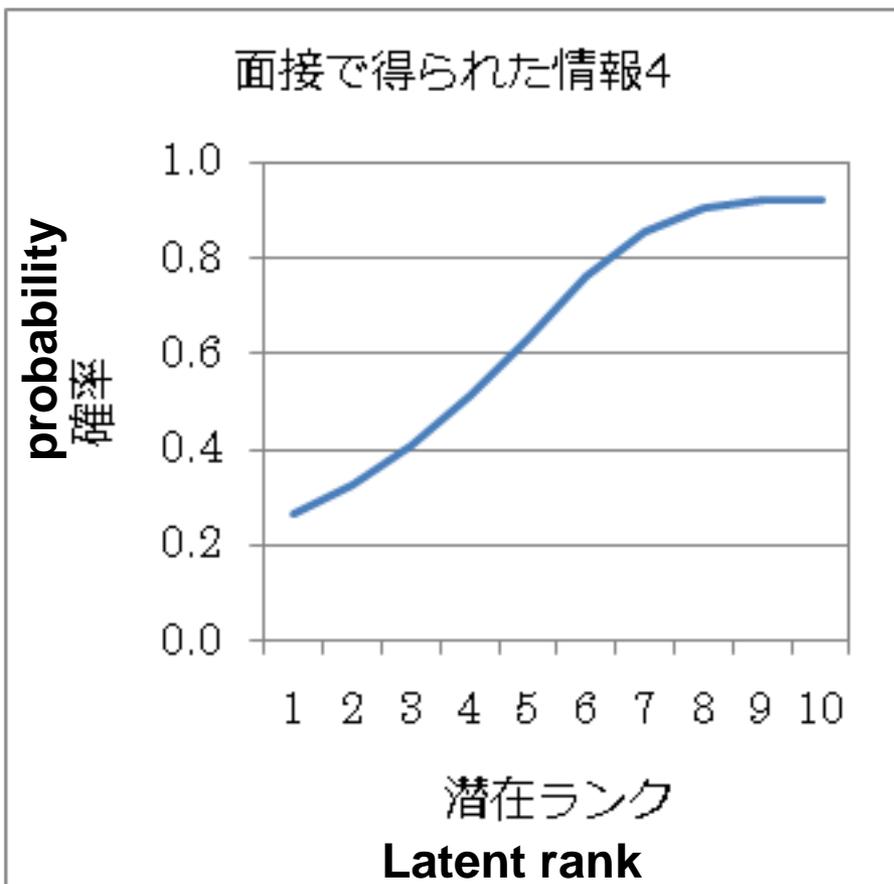
Anyway, we **overview the properties of all the items by NTT** first.

| | | "abdominal pain" | | "abdominal distension" | |
|-------------------|---|------------------|-------|------------------------|-------|
| Interview process | | average | SD | average | SD |
| 1) | Greeting | 0.981 | 0.138 | 0.934 | 0.249 |
| 2) | Selfintroduction | 0.986 | 0.120 | 0.990 | 0.101 |
| 3) | Confirmation of name | 0.976 | 0.154 | 0.974 | 0.158 |
| 4) | Listening to pt's complaints at opening | 0.966 | 0.181 | 0.974 | 0.158 |
| 5) | Appropriate, polite wording | 0.971 | 0.168 | 0.944 | 0.231 |
| 6) | Appropriate eye contact | 0.947 | 0.224 | 0.959 | 0.198 |
| 7) | Appropriate posture and attitude | 0.971 | 0.168 | 0.969 | 0.173 |
| 8) | Sympathetic words and attitude | 0.760 | 0.428 | 0.878 | 0.329 |
| 9) | Appropriate summary and confirmation | 0.933 | 0.251 | 0.918 | 0.275 |
| 10) | Ask whether anything has been misheard or for | 0.897 | 0.279 | 0.765 | 0.485 |

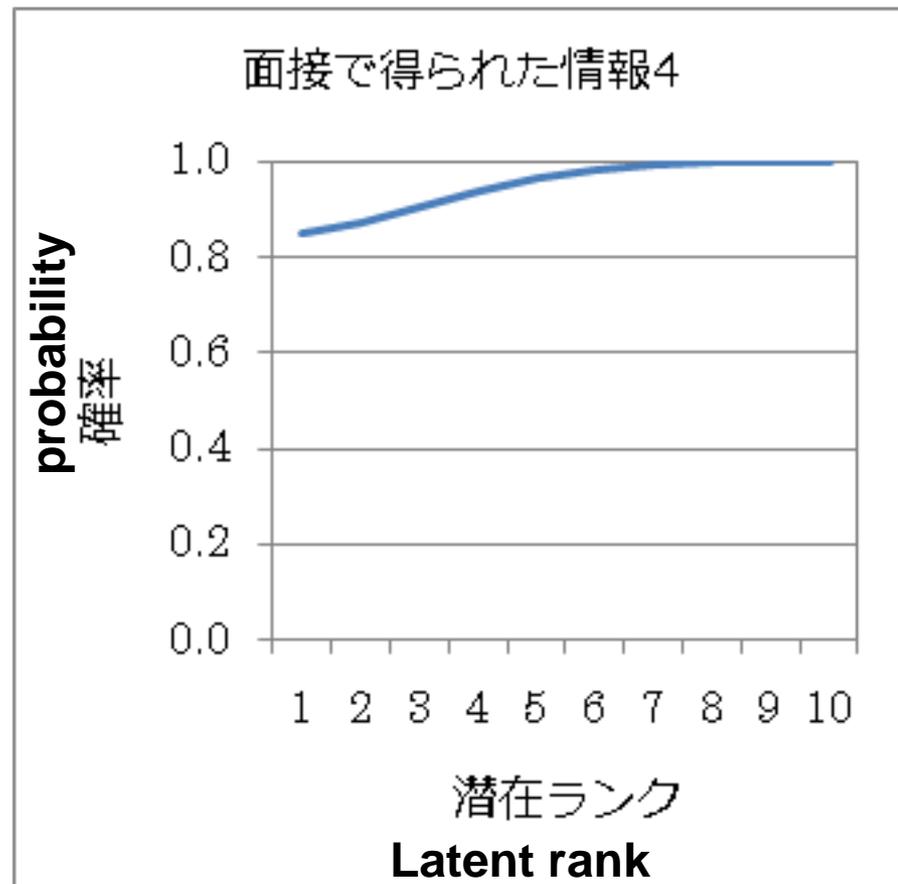
Ask for the level of the symptom

| Interview information | | | | | |
|-----------------------|--|-------|-------|-------|-------|
| 11) | Ask for the location of the symptom | 0.981 | 0.138 | 0.964 | 0.186 |
| 12) | Ask for the properties of the symptom | 0.957 | 0.204 | 0.990 | 0.101 |
| 13) | Ask for continuations of the symptom | 0.841 | 0.366 | 0.903 | 0.297 |
| 14) | Ask for the level of the symptom | 0.654 | 0.477 | 0.954 | 0.210 |
| 15) | Ask for the progress of the symptom | 1.000 | 0.000 | 0.985 | 0.123 |
| 16) | Ask for situations when the symptom occurs | 0.856 | 0.352 | 0.765 | 0.425 |
| 17) | Ask for factors in which it betters or worsens | 0.716 | 0.452 | 0.796 | 0.404 |
| 18) | Ask for accompanying symptoms | 0.981 | 0.138 | 0.939 | 0.240 |
| 19) | Ask the patient's response to the symptom | 0.928 | 0.259 | 0.816 | 0.388 |
| 20) | Ask the condition, and symptoms of the whole body. | 0.889 | 0.314 | 0.872 | 0.334 |
| 21) | Ask for the past history | 0.798 | 0.402 | 0.837 | 0.371 |
| 22) | Ask for the family history | 0.952 | 0.214 | 0.974 | 0.158 |
| 23) | Ask for extra information | 0.904 | 0.296 | 0.786 | 0.411 |
| 24) | Ask for the social situation | 0.851 | 0.357 | 0.806 | 0.396 |
| 25) | Ask for the pt's interpretation model | 0.875 | 0.332 | 0.745 | 0.437 |
| | Rating scale | 4.740 | 0.652 | 4.816 | 0.904 |

Ask for the level of the symptom



Abdominal pain



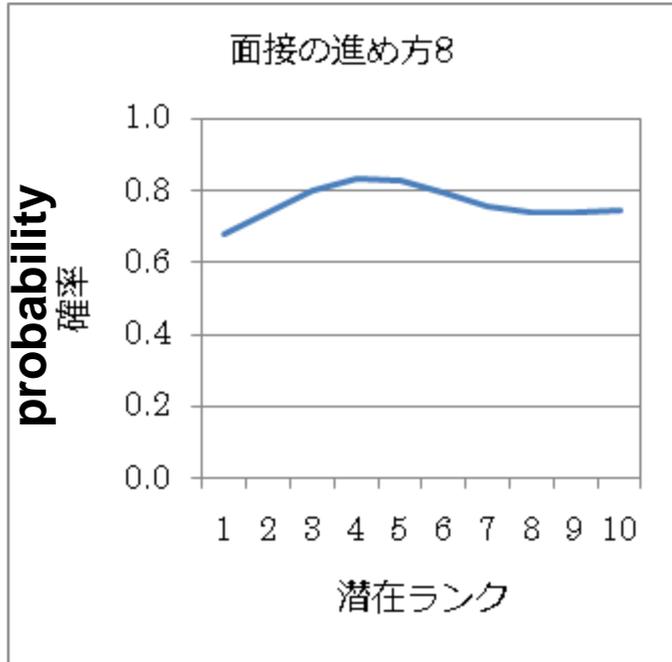
Abdominal distension

| Interview process | | "abdominal pain" | | "abdominal distension" | |
|-------------------|---|------------------|-------|------------------------|-------|
| | | average | SD | average | SD |
| 1) | Greeting | 0.981 | 0.138 | 0.934 | 0.249 |
| 2) | Selfintroduction | 0.986 | 0.120 | 0.990 | 0.101 |
| 3) | Confirmation of name | 0.976 | 0.154 | 0.974 | 0.158 |
| 4) | Listening to pt's complaints at opening | 0.966 | 0.181 | 0.974 | 0.158 |
| 5) | Appropriate, polite wording | 0.971 | 0.168 | 0.944 | 0.231 |
| 6) | Appropriate eye contact | 0.947 | 0.224 | 0.959 | 0.198 |
| 7) | Appropriate posture and attitude | 0.971 | 0.168 | 0.969 | 0.173 |
| 8) | Sympathetic words and attitude | 0.760 | 0.428 | 0.878 | 0.329 |
| 9) | Appropriate summary and confirmation | 0.933 | 0.251 | 0.918 | 0.275 |
| 10) | Ask whether anything has been misheard or for | 0.827 | 0.379 | 0.765 | 0.425 |

Sympathetic words and attitude

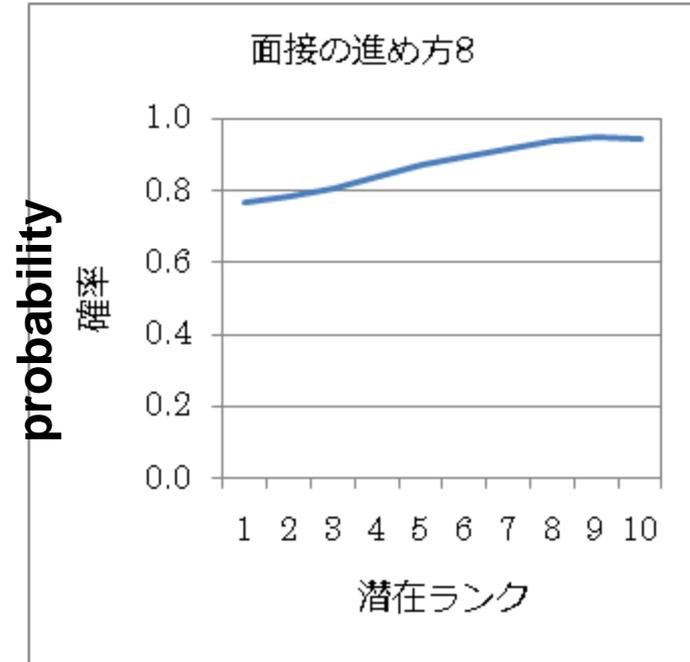
| Interview information | | | | | |
|-----------------------|--|-------|-------|-------|-------|
| 11) | Ask for the location of the symptom | 0.981 | 0.138 | 0.964 | 0.186 |
| 12) | Ask for the properties of the symptom | 0.957 | 0.204 | 0.990 | 0.101 |
| 13) | Ask for continuations of the symptom | 0.841 | 0.366 | 0.903 | 0.297 |
| 14) | Ask for the level of the symptom | 0.654 | 0.477 | 0.954 | 0.210 |
| 15) | Ask for the progress of the symptom | 1.000 | 0.000 | 0.985 | 0.123 |
| 16) | Ask for situations when the symptom occurs | 0.856 | 0.352 | 0.765 | 0.425 |
| 17) | Ask for factors in which it betters or worsens | 0.716 | 0.452 | 0.796 | 0.404 |
| 18) | Ask for accompanying symptoms | 0.981 | 0.138 | 0.939 | 0.240 |
| 19) | Ask the patient's response to the symptom | 0.928 | 0.259 | 0.816 | 0.388 |
| 20) | Ask the condition, and symptoms of the whole body. | 0.889 | 0.314 | 0.872 | 0.334 |
| 21) | Ask for the past history | 0.798 | 0.402 | 0.837 | 0.371 |
| 22) | Ask for the family history | 0.952 | 0.214 | 0.974 | 0.158 |
| 23) | Ask for extra information | 0.904 | 0.296 | 0.786 | 0.411 |
| 24) | Ask for the social situation | 0.851 | 0.357 | 0.806 | 0.396 |
| 25) | Ask for the pt's interpretation model | 0.875 | 0.332 | 0.745 | 0.437 |
| | Rating scale | 4.740 | 0.652 | 4.816 | 0.904 |

Sympathetic words and attitude



Latent rank

Abdominal pain



Latent rank

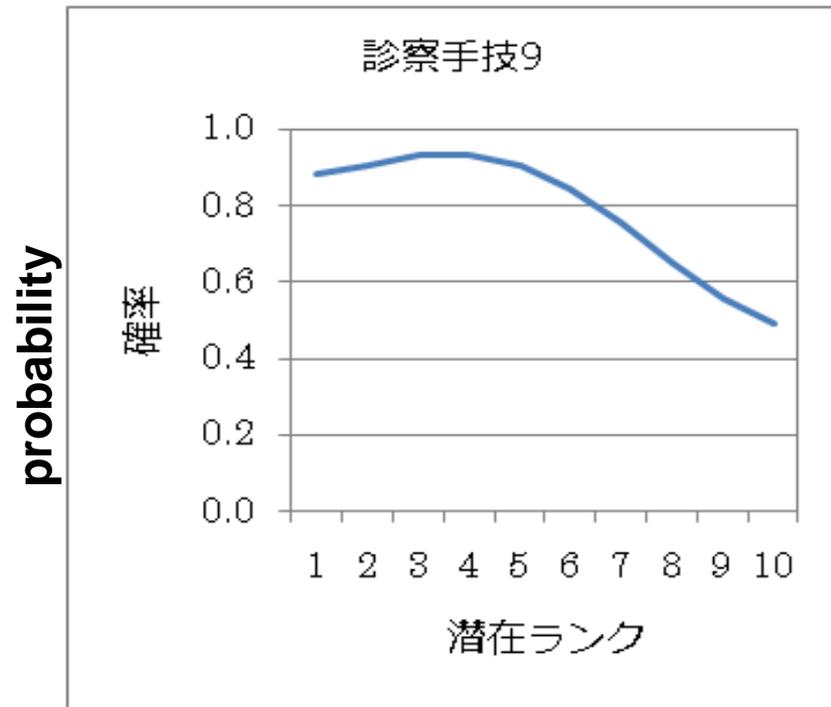
Abdominal distension

| Scenario1 "abdominal pain" | | | | Scenario2 "abdominal distension" | | | |
|----------------------------|---|---------|-------|----------------------------------|------------------------|---------|-------|
| Consideration to patient | | average | SD | | | average | SD |
| 1) | Ask for physical exam and get agreement | 0.913 | 0.282 | 1) | Same as scenario 1-1) | 0.918 | 0.275 |
| 2) | Imposing voice in each examination | 1.000 | 0.000 | 2) | Same as scenario 1-2) | 0.985 | 0.123 |
| 3) | Comprehensible imposing of voice | 0.981 | 0.138 | 3) | Same as scenario 1-3) | 0.985 | 0.123 |
| 4) | Consideration to pain | 0.990 | 0.098 | 4) | Same as scenario 1-4) | 0.980 | 0.142 |
| 5) | Warming of the stethoscope | 0.942 | 0.234 | 5) | Same as scenario 1-5) | 0.964 | 0.186 |
| Examination techniques | | | | | | | |
| 6) | Was the abdomen exposed enough? | 0.986 | 0.120 | 6) | Same as scenario 1-6) | 0.918 | 0.275 |
| 7) | Was sexuality considered? | 0.976 | 0.154 | 7) | Same as scenario 1-7) | 0.969 | 0.173 |
| 8) | Examine in order of inspection → auscultating → percussion → palpation? | 0.938 | 0.243 | 8) | Same as scenario 1-8) | 0.898 | 0.303 |
| <Inspection> | | | | <Inspection> | | | |
| 9) | Inspection of important symptoms and expression of the findings | 0.928 | 0.259 | 9) | Same as scenario 1-9) | 0.811 | 0.392 |
| <Auscultation> | | | | <Auscultation> | | | |
| 10) | Auscultate the bowel sounds and express the findings | 0.904 | 0.296 | 10) | Same as scenario 1-10) | 0.954 | 0.210 |
| 11) | Examine the pitching sounds properly and express the findings | 0.861 | 0.347 | 11) | Same as scenario 1-11) | 0.878 | 0.329 |
| <Percussion> | | | | <Percussion> | | | |
| 12) | Accurate percussion | 0.962 | 0.193 | 12) | Same as scenario 1-12) | 0.974 | 0.158 |
| 13) | Proper percussion on the the whole abdomen and confirmation of the painful area | 0.861 | 0.347 | 13) | Same as scenario 1-13) | 0.888 | 0.316 |

Were the lower limbs bent

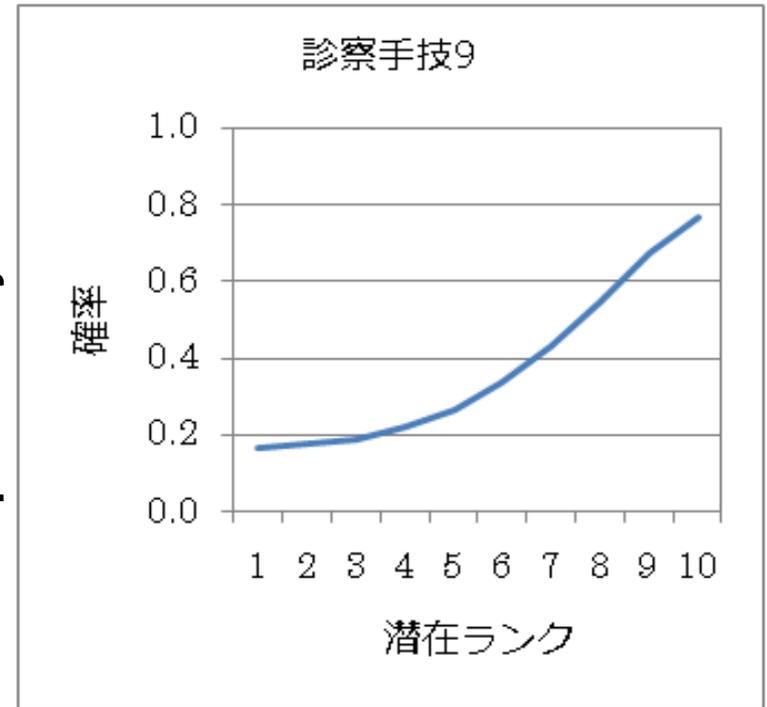
| | | | | | | | |
|--------------|---|-------|-------|--------------|--|-------|-------|
| | | | | 16) | Was spleen enlargement confirmed by accurate percussion? | 0.689 | 0.464 |
| <Palpation> | | | | <Palpation> | | | |
| 14) | Were the lower limbs bent | 0.774 | 0.419 | 17) | Same as scenario 1-14) | 0.801 | 0.400 |
| 15) | Proper superficial palpation on the whole abdomen first | 0.947 | 0.224 | 18) | Same as scenario 1-15) | 0.964 | 0.186 |
| 16) | Was the deep palpation performed correctly next? | 0.938 | 0.243 | 19) | Same as scenario 1-16) | 0.939 | 0.240 |
| 17) | Was the painful part (in the right low quadrant) palpated at the end? | 0.803 | 0.399 | 20) | Was the liver confirmed by accurate palpation? | 0.776 | 0.418 |
| 18) | Was the pain pressure point confirmed? | 0.745 | 0.437 | | | | |
| 19) | Was rebound tenderness confirmed? | 0.750 | 0.434 | | | | |
| 20) | Was muscular guarding confirmed? | 0.620 | 0.487 | | | | |
| Rating scale | | 4.808 | 0.793 | Rating scale | | 4.699 | 0.953 |

Were the lower limbs bent



Latent rank

Abdominal pain



Latent rank

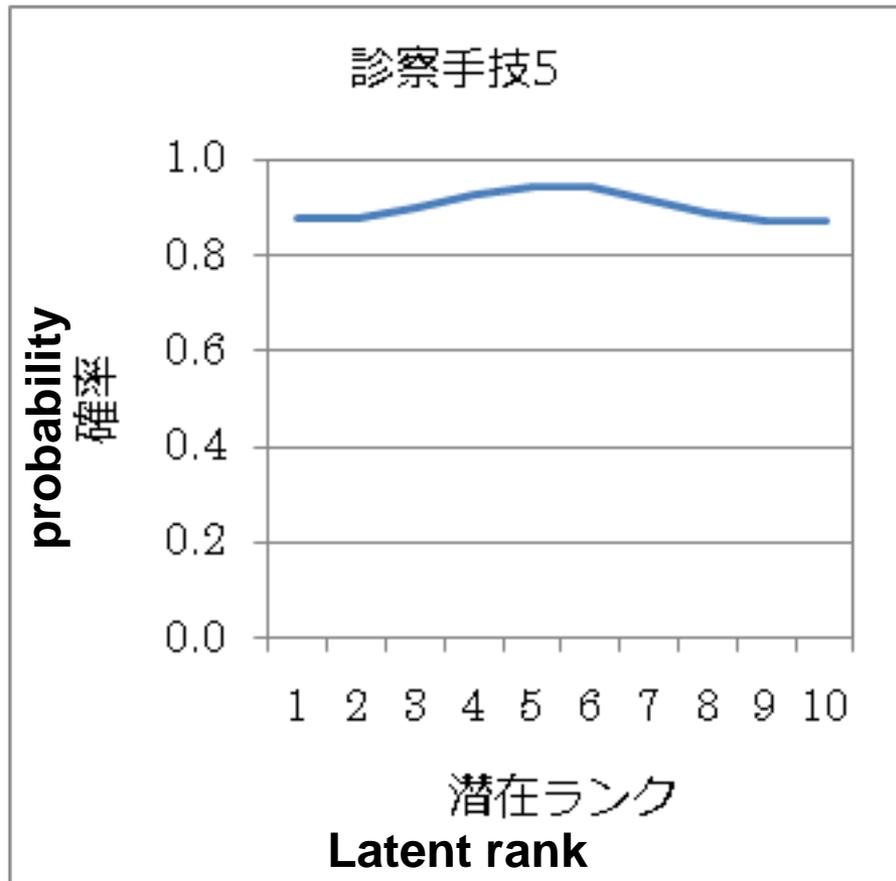
Abdominal distension

| Scenario1 " abdominal pain " | | average | SD | Scenario2 | " abdominal distension " | average | SD |
|------------------------------|---|---------|-------|----------------|--------------------------|---------|-------|
| Consideration to patient | | | | | | | |
| 1) | Ask for physical exam and get agreement | 0.913 | 0.282 | 1) | Same as scenario 1-1) | 0.918 | 0.275 |
| 2) | Imposing voice in each examination | 1.000 | 0.000 | 2) | Same as scenario 1-2) | 0.985 | 0.123 |
| 3) | Comprehensible imposing of voice | 0.981 | 0.138 | 3) | Same as scenario 1-3) | 0.985 | 0.123 |
| 4) | Consideration to pain | 0.990 | 0.098 | 4) | Same as scenario 1-4) | 0.980 | 0.142 |
| 5) | Warming of the stethoscope | 0.942 | 0.234 | 5) | Same as scenario 1-5) | 0.964 | 0.186 |
| Examination techniques | | | | | | | |
| 6) | Was the abdomen exposed enough? | 0.986 | 0.120 | 6) | Same as scenario 1-6) | 0.918 | 0.275 |
| 7) | Was sexuality considered? | 0.976 | 0.154 | 7) | Same as scenario 1-7) | 0.969 | 0.173 |
| 8) | Examine in order of inspection → auscultating → percussion → palpation? | 0.938 | 0.243 | 8) | Same as scenario 1-8) | 0.898 | 0.303 |
| <Inspection> | | | | <Inspection> | | | |
| 9) | Inspection of important symptoms and expression of the findings | 0.928 | 0.259 | 9) | Same as scenario 1-9) | 0.811 | 0.392 |
| <Auscultation> | | | | <Auscultation> | | | |
| 10) | Auscultate the bowel sounds and express the findings | 0.904 | 0.296 | 10) | Same as scenario 1-10) | 0.954 | 0.210 |
| 11) | Examine the pitching sounds properly and express the findings | 0.861 | 0.347 | 11) | Same as scenario 1-11) | 0.878 | 0.329 |
| <Percussion> | | | | <Percussion> | | | |
| 12) | Accurate percussion | 0.962 | 0.193 | 12) | Same as scenario 1-12) | 0.974 | 0.158 |
| 13) | Proper percussion on the the whole abdomen and | 0.861 | 0.347 | 13) | Same as scenario 1-13) | 0.888 | 0.316 |

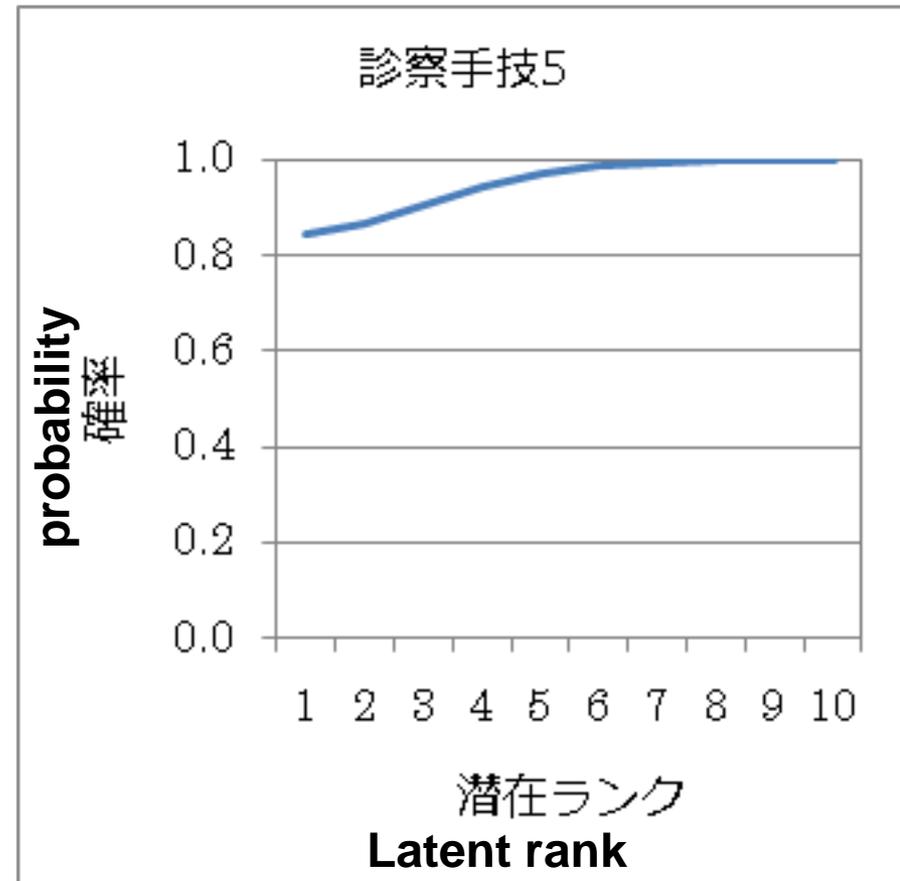
Auscultate the bowel sounds and express the findings

| | | | | | | | |
|---------------|---|-------|-------|---------------|--|-------|-------|
| <Palpitation> | | | | <Palpitation> | | | |
| 14) | Were the lower limbs bent | 0.774 | 0.419 | 17) | Same as scenario 1-14) | 0.801 | 0.400 |
| 15) | Proper superficial palpitation on the whole abdomen first | 0.947 | 0.224 | 18) | Same as scenario 1-15) | 0.964 | 0.186 |
| 16) | Was the deep palpation performed correctly next? | 0.938 | 0.243 | 19) | Same as scenario 1-16) | 0.939 | 0.240 |
| 17) | Was the painful part (in the right low quadrant) palpated at the end? | 0.803 | 0.399 | 20) | Was the liver confirmed by accurate palpation? | 0.776 | 0.418 |
| 18) | Was the pain pressure point confirmed? | 0.745 | 0.437 | | | | |
| 19) | Was rebound tenderness confirmed? | 0.750 | 0.434 | | | | |
| 20) | Was muscular guarding confirmed? | 0.620 | 0.487 | | | | |
| Rating scale | | 4.808 | 0.793 | Rating scale | | 4.699 | 0.953 |

Auscultate the bowel sounds and express the findings



Abdominal pain



Abdominal distension

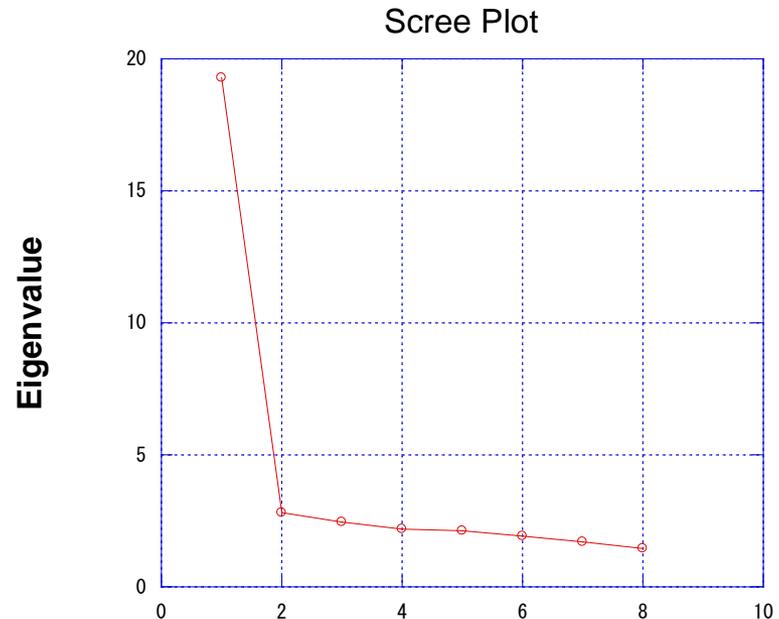
The Categorical Factor Analyses were performed, **Eigen-values** and **factor loading values** were calculated using tetrachoric correlation matrix.

The **11 items** and the **23 items** were chosen from “abdominal pain” and “abdominal distension”.

The first Eigen-values were much larger than those of the second group.



Scree Plot of Abdominal pain



Scree Plot of Abdominal distension

Factor loadings for "abdominal pain" and "abdominal distension"

| Scenario 1 "abdominal pain" | | | Scenario 2 "abdominal distension" | | |
|-----------------------------|----------|-----------------------------------|-----------------------------------|----------|-----------------------------------|
| | 11 Items | 11Items +2Rating +2Subtotal | | 23 Items | 11Items +2Rating +2Subtotal |
| Interview process 6 | 0.79 | 0.76 | Interview process 1 | 0.685 | 0.743 |
| Interview process 7 | 0.688 | 0.635 | Interview process 8 | 0.567 | 0.614 |
| Interview information 4 | 0.443 | 0.557 | Interview process 9 | 0.44 | 0.472 |
| Interview information 14 | 0.395 | 0.542 | Interview process 10 | 0.392 | 0.525 |
| Rating scale | | 0.724 | Interview information 3 | 0.545 | 0.542 |
| Subtotal | | 0.644 | Interview information 4 | 0.696 | 0.648 |
| | | | Interview information 6 | 0.445 | 0.615 |
| Examination technique 1 | 0.733 | 0.606 | Interview information 7 | 0.373 | 0.516 |
| Examination technique 4 | 0.575 | 0.562 | Interview information 8 | 0.642 | 0.625 |
| Examination technique 7 | 0.569 | 0.624 | Interview information 10 | 0.393 | 0.396 |
| Examination technique 10 | 0.819 | 0.761 | Interview information 11 | 0.308 | 0.366 |
| Examination technique 11 | 0.803 | 0.768 | Rating scale | | 0.674 |
| Examination technique 13 | 0.544 | 0.518 | Subtotal | | 0.8 |
| Examination technique 14 | 0.358 | 0.499 | | | |
| Rating scale | | 0.832 | Consideration to patient 3 | 0.534 | 0.587 |
| Subtotal | | 0.76 | Examination technique 3 | 0.602 | 0.604 |
| | | | Examination technique 4 | 0.314 | 0.367 |
| | | | Examination technique 5 | 0.741 | 0.673 |
| | | | Examination technique 6 | 0.421 | 0.368 |
| | | | Examination technique 8 | 0.616 | 0.655 |
| | | | Examination technique 9 | 0.469 | 0.589 |
| | | | Examination technique 10 | 0.514 | 0.613 |
| | | | Examination technique 11 | 0.421 | 0.482 |
| | | | Examination technique 12 | 0.434 | 0.512 |
| | | | Examination technique 14 | 0.581 | 0.482 |
| | | | Examination technique 15 | 0.353 | 0.35 |
| | | | Rating scale | | 0.804 |
| | | | Subtotal | | 0.883 |

The 11 items for "abdominal pain" and the 23 items for "abdominal distension" of which factor loading values exceeded 0.3, were chosen according to Lord and Saida's criteria.

Methods of IRT analysis.

The IRT was applied to the selected assessment items.

The values of these difficulty and discrimination parameters were calculated for these items.

We used **2-parameter-model** IRT calculated by **M-PLUS** with the **weighted least square method**.

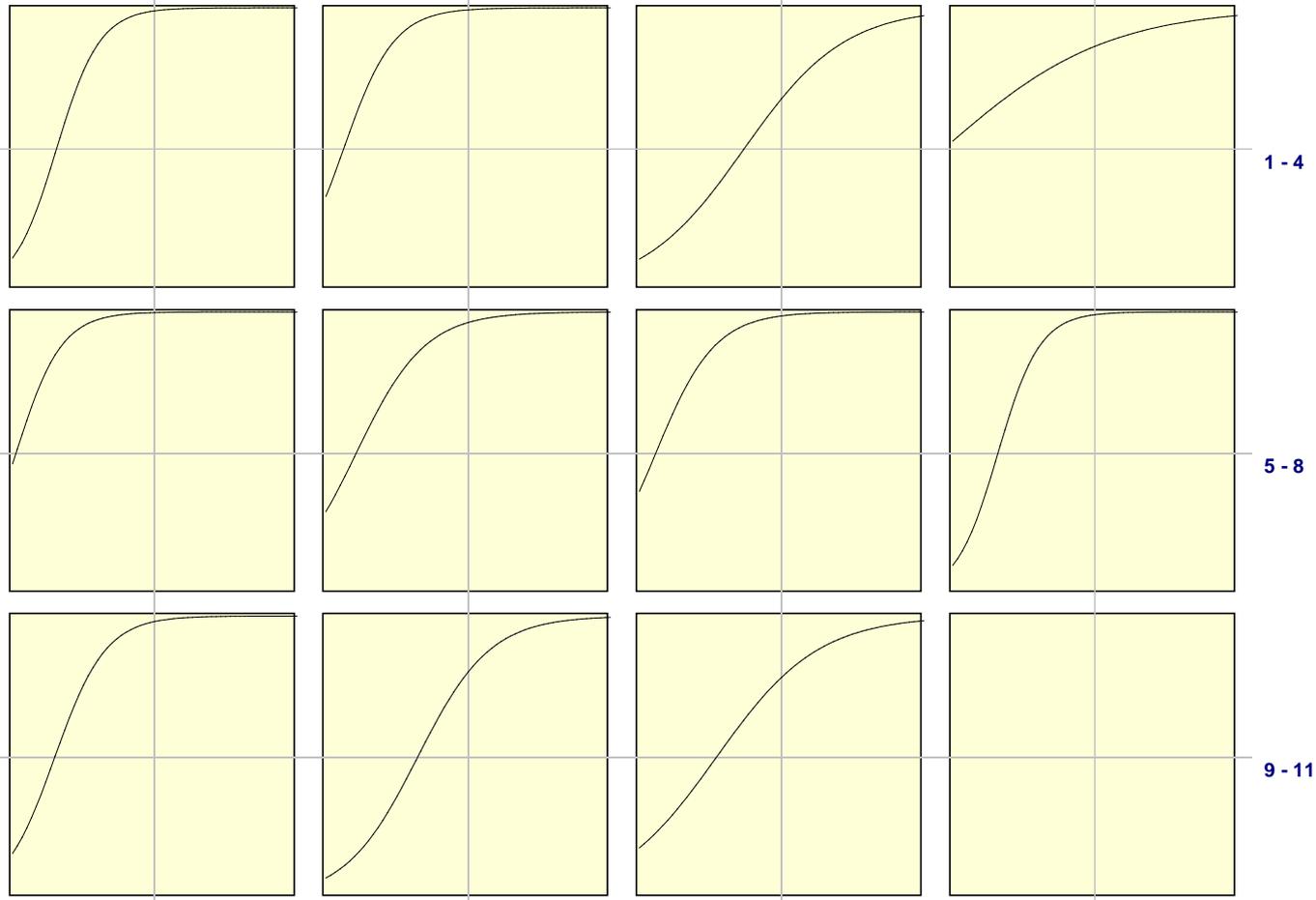
The examinee's **abilities** were calculated by **BILOG-MG** with **Bayes (EAP)** method.

Estimate and P value of Discrimination and Difficulty for abdominal pain and abdominal distension

| Table 3 "Estimated and P value of Discrimination and Difficulty for "abdominal pain" and "abdominal distension" | | | | | | | | |
|---|----------------|-------|-----------|---------|-----------------------|-------|-----------|---------|
| Scenario1"abdominal pain" | Discrimination | | | | Two-Tailed Difficulty | | | |
| | Estimate | S.E.. | Est./S.E. | P-Value | Estimate | S.E.. | Est./S.E. | P-Value |
| Interview process 6 | 1.395 | 0.534 | 2.611 | 0.009 | -1.99 | 0.326 | -6.101 | 0 |
| Interview process 7 | 0.901 | 0.353 | 2.554 | 0.011 | -2.835 | 0.674 | -4.204 | 0 |
| Interview information 4 | 0.526 | 0.14 | 3.763 | 0 | -0.85 | 0.257 | -3.306 | 0.001 |
| Interview information 14 | 0.405 | 0.157 | 2.584 | 0.01 | -4.197 | 1.381 | -3.038 | 0.002 |
| Examination technique 1 | 1.079 | 0.261 | 4.139 | 0 | -2.98 | 0.434 | -6.86 | 0 |
| Examination technique 4 | 0.784 | 0.28 | 2.795 | 0.005 | -2.367 | 0.579 | -4.088 | 0 |
| Examination technique 7 | 0.858 | 0.339 | 2.526 | 0.012 | -2.717 | 0.703 | -3.865 | 0 |
| Examination technique 10 | 1.89 | 0.869 | 2.175 | 0.03 | -1.83 | 0.253 | -7.242 | 0 |
| Examination technique 11 | 1.372 | 0.665 | 2.061 | 0.039 | -1.899 | 0.372 | -5.108 | 0 |
| Examination technique 13 | 0.721 | 0.174 | 4.143 | 0 | -1.128 | 0.238 | -4.736 | 0 |
| Examination technique 14 | 0.387 | 0.14 | 2.775 | 0.006 | -1.868 | 0.637 | -2.931 | 0.003 |
| average | 0.938 | | | | -2.242 | | | |
| standard deviation | 0.465 | | | | 0.929 | | | |
| | | | | | | | | |
| Scenario2"abdominal distension" | Discrimination | | | | Two-Tailed Difficulty | | | |
| | Estimate | S.E.. | Est./S.E. | P-Value | Estimate | S.E.. | Est./S.E. | P-Value |
| Interview process 1 | 1.231 | 0.401 | 3.066 | 0.002 | -1.937 | 0.294 | -6.585 | 0 |
| Interview process 8 | 0.648 | 0.164 | 3.955 | 0 | -2.139 | 0.433 | -4.937 | 0 |
| Interview process 9 | 0.564 | 0.266 | 2.125 | 0.034 | -2.837 | 1.039 | -2.73 | 0.006 |
| Interview process 10 | 0.471 | 0.11 | 4.295 | 0 | -1.698 | 0.391 | -4.345 | 0 |
| Interview information 3 | 0.884 | 0.193 | 4.578 | 0 | -1.961 | 0.296 | -6.628 | 0 |
| Interview information 4 | 1.447 | 0.685 | 2.113 | 0.035 | -2.049 | 0.345 | -5.944 | 0 |
| Interview information 6 | 0.852 | 0.125 | 6.791 | 0 | -1.115 | 0.176 | -6.336 | 0 |
| Interview information 7 | 0.747 | 0.131 | 5.72 | 0 | -1.382 | 0.23 | -6.014 | 0 |
| Interview information 8 | 0.858 | 0.344 | 2.496 | 0.013 | -2.372 | 0.585 | -4.051 | 0 |
| Interview information 10 | 0.461 | 0.167 | 2.766 | 0.006 | -2.72 | 0.855 | -3.182 | 0.001 |
| Interview information 11 | 0.333 | 0.15 | 2.222 | 0.026 | -3.105 | 1.298 | -2.392 | 0.017 |
| Consideration to patient 3 | 0.663 | 0.241 | 2.747 | 0.006 | -2.523 | 0.67 | -3.763 | 0 |
| Examination technique 3 | 0.785 | 0.211 | 3.718 | 0 | -2.057 | 0.386 | -5.325 | 0 |
| Examination technique 4 | 0.329 | 0.141 | 2.327 | 0.02 | -2.824 | 1.139 | -2.478 | 0.013 |
| Examination technique 5 | 1.314 | 0.597 | 2.201 | 0.028 | -2.118 | 0.381 | -5.558 | 0 |
| Examination technique 6 | 0.502 | 0.18 | 2.791 | 0.005 | -2.593 | 0.782 | -3.314 | 0.001 |
| Examination technique 8 | 0.828 | 0.176 | 4.704 | 0 | -1.904 | 0.295 | -6.452 | 0 |
| Examination technique 9 | 0.505 | 0.108 | 4.672 | 0 | 0.515 | 0.224 | 2.302 | 0.021 |
| Examination technique 10 | 0.703 | 0.14 | 5.025 | 0 | -1.779 | 0.297 | -5.986 | 0 |
| Examination technique 11 | 0.475 | 0.1 | 4.755 | 0 | -1.148 | 0.286 | -4.018 | 0 |
| Examination technique 12 | 0.524 | 0.126 | 4.146 | 0 | -1.82 | 0.404 | -4.501 | 0 |
| Examination technique 14 | 0.716 | 0.306 | 2.338 | 0.019 | -2.654 | 0.783 | -3.388 | 0.001 |
| Examination technique 15 | 0.378 | 0.118 | 3.197 | 0.001 | -2.14 | 0.646 | -3.315 | 0.001 |
| average | 0.705 | | | | -2.016 | | | |
| standard deviation | 0.301 | | | | 0.762 | | | |

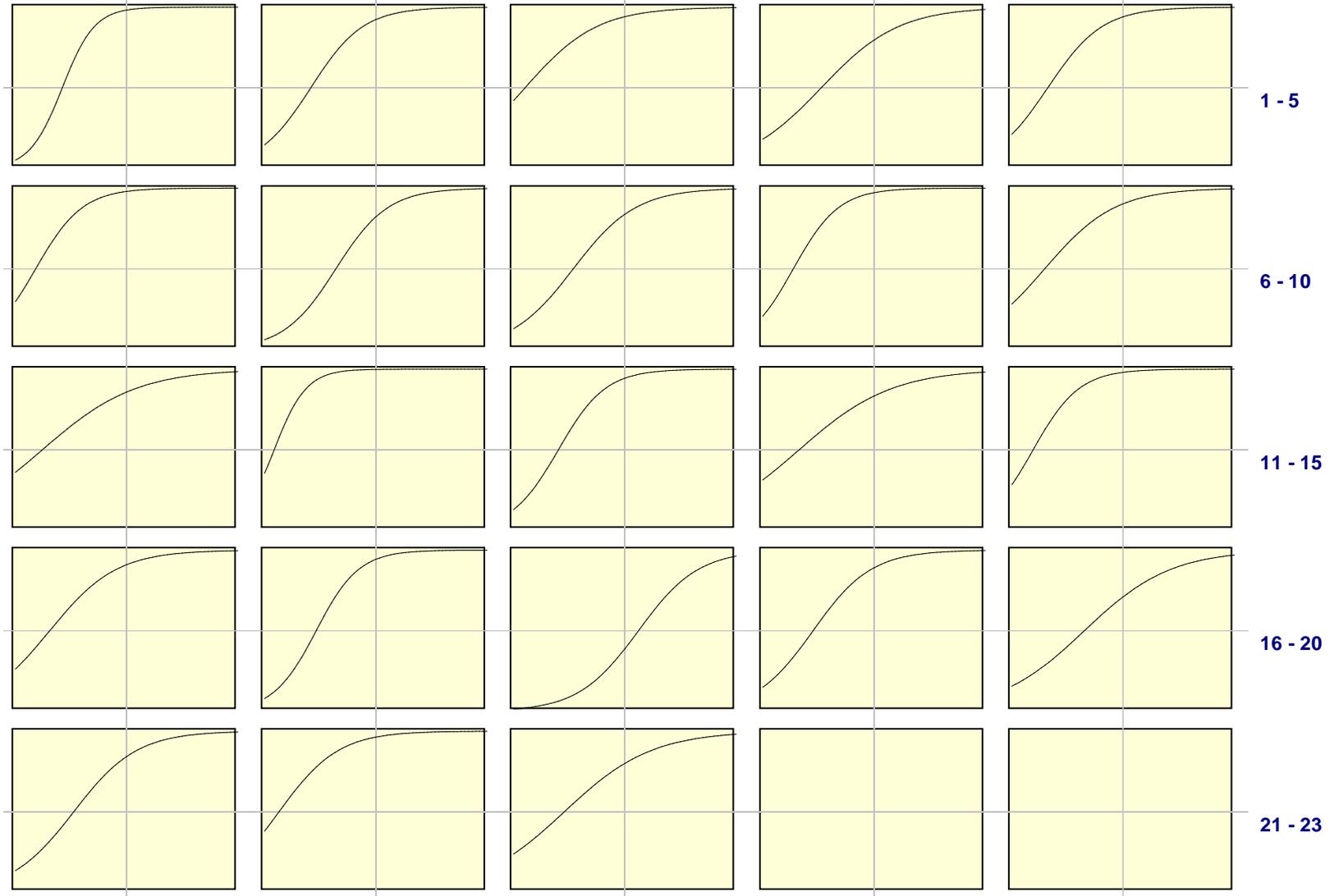
ICC for 11 items of abdominal pain

Matrix Plot of Item Characteristic Curves



ICC for 23 items for abdominal distension

Matrix Plot of Item Characteristic Curves



Parameters estimated by MCMC

Abdominal pain station

| | Mean | SD | | |
|-------|----------|---------|-------|-------------------|
| a[1] | 1.096333 | 0.27765 | b[1] | -2.319168 0.38018 |
| a[2] | 1.153846 | 0.29455 | b[2] | -2.741558 0.44706 |
| a[3] | 0.373296 | 0.12275 | b[3] | -1.290439 0.48914 |
| a[4] | 0.350467 | 0.08648 | b[4] | -3.173517 0.70126 |
| a[5] | 0.577399 | 0.14078 | b[5] | -3.160971 0.59866 |
| a[6] | 0.825345 | 0.20725 | b[6] | -2.594877 0.49384 |
| a[7] | 0.997992 | 0.26139 | b[7] | -2.631976 0.47030 |
| a[8] | 1.234399 | 0.31333 | b[8] | -2.181446 0.31529 |
| a[9] | 1.148370 | 0.28005 | b[9] | -1.976976 0.28394 |
| a[10] | 0.775217 | 0.20800 | b[10] | -1.127125 0.26087 |
| a[11] | 0.541580 | 0.13954 | b[11] | -1.615191 0.37767 |

Parameters estimated by MCMC

Abdominal distension station

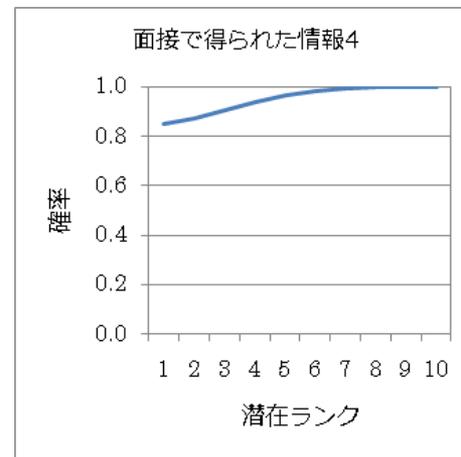
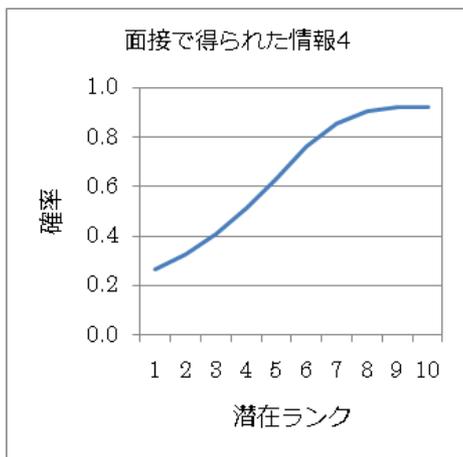
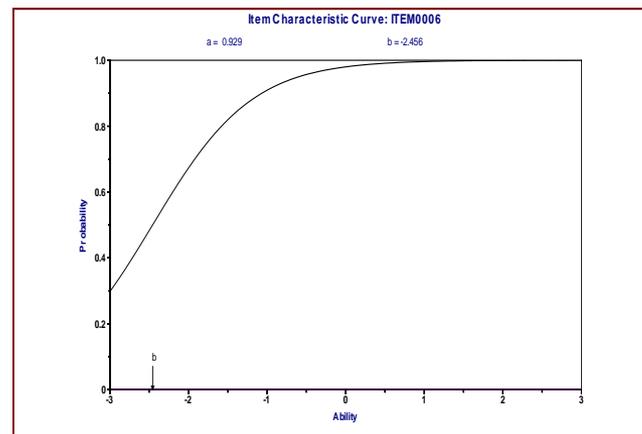
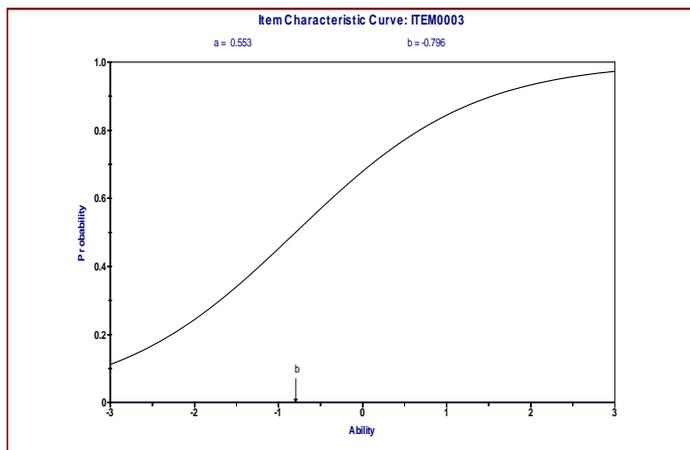
| | Mean | SD | - | | |
|-------|----------|--------|---|-------|------------------|
| a[1] | 1.134116 | 0.3324 | | b[1] | -1.957563 0.4246 |
| a[2] | 0.728898 | 0.2182 | | b[2] | -1.969713 0.5080 |
| a[3] | 0.541417 | 0.1418 | | b[3] | -2.990348 0.7047 |
| a[4] | 0.458977 | 0.1405 | | b[4] | -1.637180 0.5476 |
| a[5] | 0.721120 | 0.2115 | | b[5] | -2.245857 0.5883 |
| a[6] | 0.842489 | 0.2459 | | b[6] | -2.714691 0.6242 |
| a[7] | 0.732104 | 0.2183 | | b[7] | -1.159403 0.3405 |
| a[8] | 0.575587 | 0.1897 | | b[8] | -1.677140 0.5793 |
| a[9] | 0.807309 | 0.2282 | | b[9] | -2.368200 0.5575 |
| a[10] | 0.521406 | 0.1503 | | b[10] | -2.477535 0.6283 |
| a[11] | 0.379173 | 0.1030 | | b[11] | -2.728008 0.6935 |
| a[12] | 1.206744 | 0.3418 | | b[12] | -3.051057 0.6393 |
| a[13] | 0.776014 | 0.2268 | | b[13] | -1.983169 0.4914 |
| a[14] | 0.375527 | 0.1095 | | b[14] | -2.480604 0.6766 |
| a[15] | 0.874504 | 0.2512 | | b[15] | -2.635444 0.5972 |
| a[16] | 0.546648 | 0.1558 | | b[16] | -2.453949 0.6138 |
| a[17] | 0.888031 | 0.2580 | | b[17] | -1.809220 0.4294 |
| a[18] | 0.641890 | 0.1854 | | b[18] | 0.550426 0.2152 |
| a[19] | 0.631000 | 0.1884 | | b[19] | -1.912281 0.5208 |
| a[20] | 0.346777 | 0.1238 | | b[20] | -1.512952 0.6127 |
| a[21] | 0.538801 | 0.1628 | | b[21] | -1.782483 0.5300 |
| a[22] | 0.669474 | 0.1911 | | b[22] | -2.877118 0.6786 |
| a[23] | 0.371681 | 0.1142 | | b[23] | -2.108927 0.6494 |

| Interview process | | "abdominal pain" | | "abdominal distension" | |
|-------------------|---|------------------|-------|------------------------|-------|
| | | average | SD | average | SD |
| 1) | Greeting | 0.981 | 0.138 | 0.934 | 0.249 |
| 2) | Selfintroduction | 0.986 | 0.120 | 0.990 | 0.101 |
| 3) | Confirmation of name | 0.976 | 0.154 | 0.974 | 0.158 |
| 4) | Listening to pt's complaints at opening | 0.966 | 0.181 | 0.974 | 0.158 |
| 5) | Appropriate, polite wording | 0.971 | 0.168 | 0.944 | 0.231 |
| 6) | Appropriate eye contact | 0.947 | 0.224 | 0.959 | 0.198 |
| 7) | Appropriate posture and attitude | 0.971 | 0.168 | 0.969 | 0.173 |
| 8) | Sympathetic words and attitude | 0.760 | 0.428 | 0.878 | 0.329 |
| 9) | Appropriate summary and confirmation | 0.933 | 0.251 | 0.918 | 0.275 |
| 10) | Ask whether anything has been misheard or for | 0.827 | 0.379 | 0.765 | 0.425 |

Ask for the level of the symptom

| Interview information | | | | | |
|-----------------------|--|-------|-------|-------|-------|
| 11) | Ask for the location of the symptom | 0.981 | 0.138 | 0.964 | 0.186 |
| 12) | Ask for the properties of the symptom | 0.957 | 0.204 | 0.990 | 0.101 |
| 13) | Ask for continuations of the symptom | 0.841 | 0.366 | 0.903 | 0.297 |
| 14) | Ask for the level of the symptom | 0.654 | 0.477 | 0.954 | 0.210 |
| 15) | Ask for the progress of the symptom | 1.000 | 0.000 | 0.985 | 0.123 |
| 16) | Ask for situations when the symptom occurs | 0.856 | 0.352 | 0.765 | 0.425 |
| 17) | Ask for factors in which it betters or worsens | 0.716 | 0.452 | 0.796 | 0.404 |
| 18) | Ask for accompanying symptoms | 0.981 | 0.138 | 0.939 | 0.240 |
| 19) | Ask the patient's response to the symptom | 0.928 | 0.259 | 0.816 | 0.388 |
| 20) | Ask the condition, and symptoms of the whole body. | 0.889 | 0.314 | 0.872 | 0.334 |
| 21) | Ask for the past history | 0.798 | 0.402 | 0.837 | 0.371 |
| 22) | Ask for the family history | 0.952 | 0.214 | 0.974 | 0.158 |
| 23) | Ask for extra information | 0.904 | 0.296 | 0.786 | 0.411 |
| 24) | Ask for the social situation | 0.851 | 0.357 | 0.806 | 0.396 |
| 25) | Ask for the pt's interpretation model | 0.875 | 0.332 | 0.745 | 0.437 |
| | Rating scale | 4.740 | 0.652 | 4.816 | 0.904 |

Ask for the level of the symptom



Abdominal pain

Abdominal distension

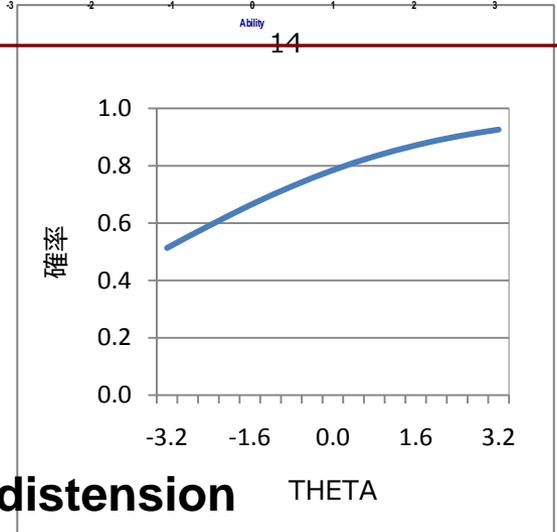
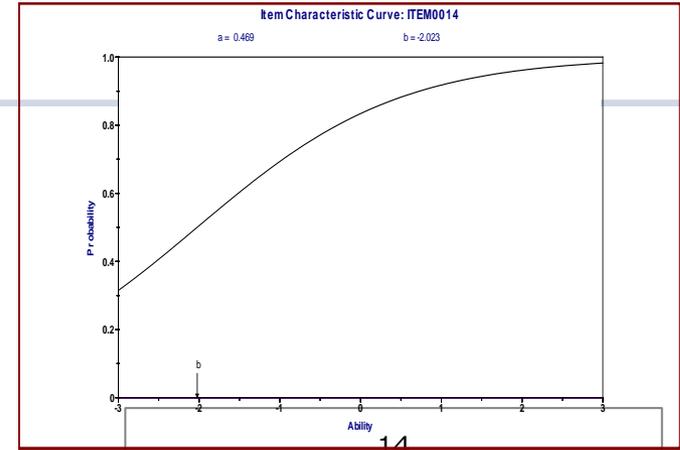
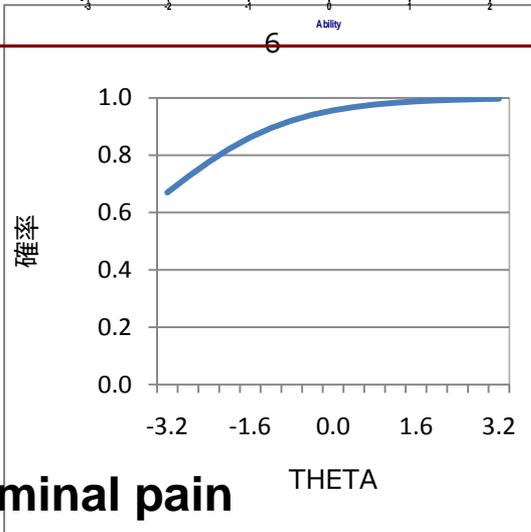
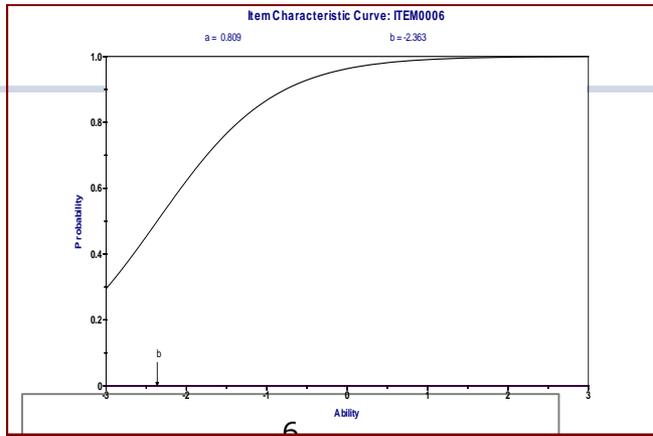
| | discrimination | difficulty |
|----------------------|----------------|------------|
| abdominal pain | 0.526 | -0.85 |
| abdominal distension | 1.447 | -2.049 |

| Scenario1 | " abdominal pain" | average | SD | Scenario2 | " abdominal distension" | average | SD |
|--------------------------|---|---------|-------|----------------|-------------------------|---------|-------|
| Consideration to patient | | | | | | | |
| 1) | Ask for physical exam and get agreement | 0.913 | 0.282 | 1) | Same as scenario 1-1) | 0.918 | 0.275 |
| 2) | Imposing voice in each examination | 1.000 | 0.000 | 2) | Same as scenario 1-2) | 0.985 | 0.123 |
| 3) | Comprehensible imposing of voice | 0.981 | 0.138 | 3) | Same as scenario 1-3) | 0.985 | 0.123 |
| 4) | Consideration to pain | 0.990 | 0.098 | 4) | Same as scenario 1-4) | 0.980 | 0.142 |
| 5) | Warming of the stethoscope | 0.942 | 0.234 | 5) | Same as scenario 1-5) | 0.964 | 0.186 |
| Examination techniques | | | | | | | |
| 6) | Was the abdomen exposed enough? | 0.986 | 0.120 | 6) | Same as scenario 1-6) | 0.918 | 0.275 |
| 7) | Was sexuality considered? | 0.976 | 0.154 | 7) | Same as scenario 1-7) | 0.969 | 0.173 |
| 8) | Examine in order of inspection → auscultating → percussion → palpation? | 0.938 | 0.243 | 8) | Same as scenario 1-8) | 0.898 | 0.303 |
| <Inspection> | | | | <Inspection> | | | |
| 9) | Inspection of important symptoms and expression of the findings | 0.928 | 0.259 | 9) | Same as scenario 1-9) | 0.811 | 0.392 |
| <Auscultation> | | | | <Auscultation> | | | |
| 10) | Auscultate the bowel sounds and express the findings | 0.904 | 0.296 | 10) | Same as scenario 1-10) | 0.954 | 0.210 |
| 11) | Examine the pitching sounds properly and express the findings | 0.861 | 0.347 | 11) | Same as scenario 1-11) | 0.878 | 0.329 |
| <Percussion> | | | | <Percussion> | | | |
| 12) | Accurate percussion | 0.962 | 0.193 | 12) | Same as scenario 1-12) | 0.974 | 0.158 |
| 13) | Proper percussion on the the whole abdomen and confirmation of the painful area | 0.861 | 0.347 | 13) | Same as scenario 1-13) | 0.888 | 0.316 |

Inspection of important symptoms and expression of the findings

| | | | | | | | |
|--------------|---|-------|-------|--------------|--|-------|-------|
| 14) | abdomen first | 0.817 | 0.221 | 14) | Same as scenario 1-14) | 0.881 | 0.188 |
| 16) | Was the deep palpation performed correctly next? | 0.938 | 0.243 | 19) | Same as scenario 1-16) | 0.939 | 0.240 |
| 17) | Was the painful part (in the right low quadrant) palpated at the end? | 0.803 | 0.399 | 20) | Was the liver confirmed by accurate palpation? | 0.776 | 0.418 |
| 18) | Was the pain pressure point confirmed? | 0.745 | 0.437 | | | | |
| 19) | Was rebound tenderness confirmed? | 0.750 | 0.434 | | | | |
| 20) | Was muscular guarding confirmed? | 0.620 | 0.487 | | | | |
| Rating scale | | 4.808 | 0.793 | Rating scale | | 4.699 | 0.953 |

Inspection of important symptoms and expression of the findings

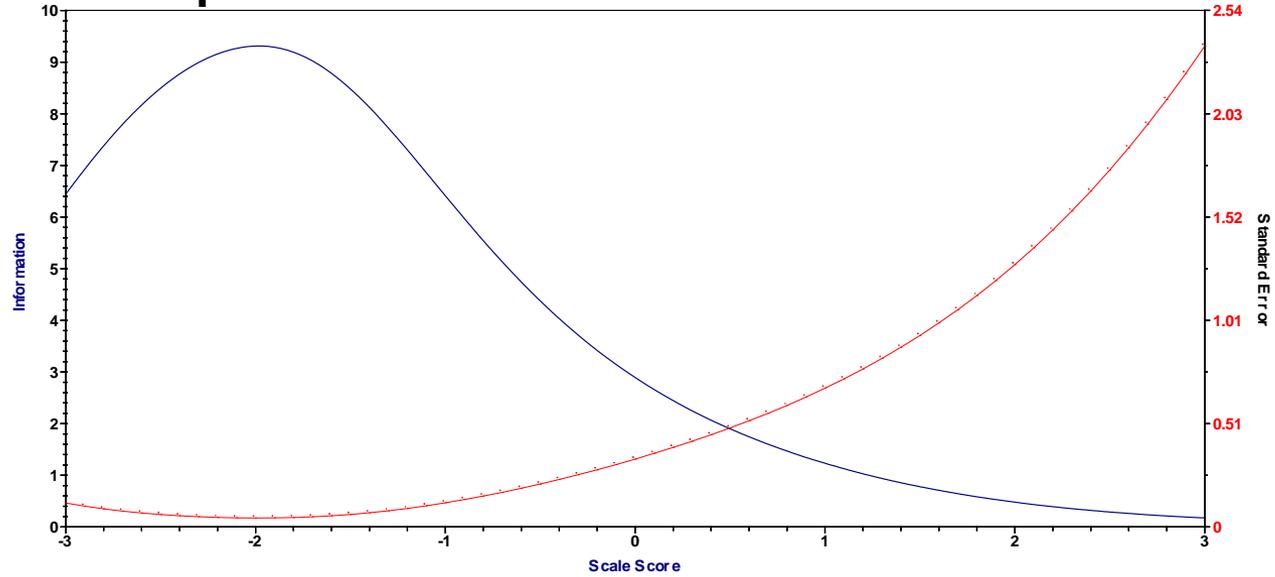


| | discrimination | difficulty |
|----------------------|----------------|------------|
| abdominal pain | 0.784 | -2.367 |
| abdominal distension | 0.329 | -2.824 |

Information Curve

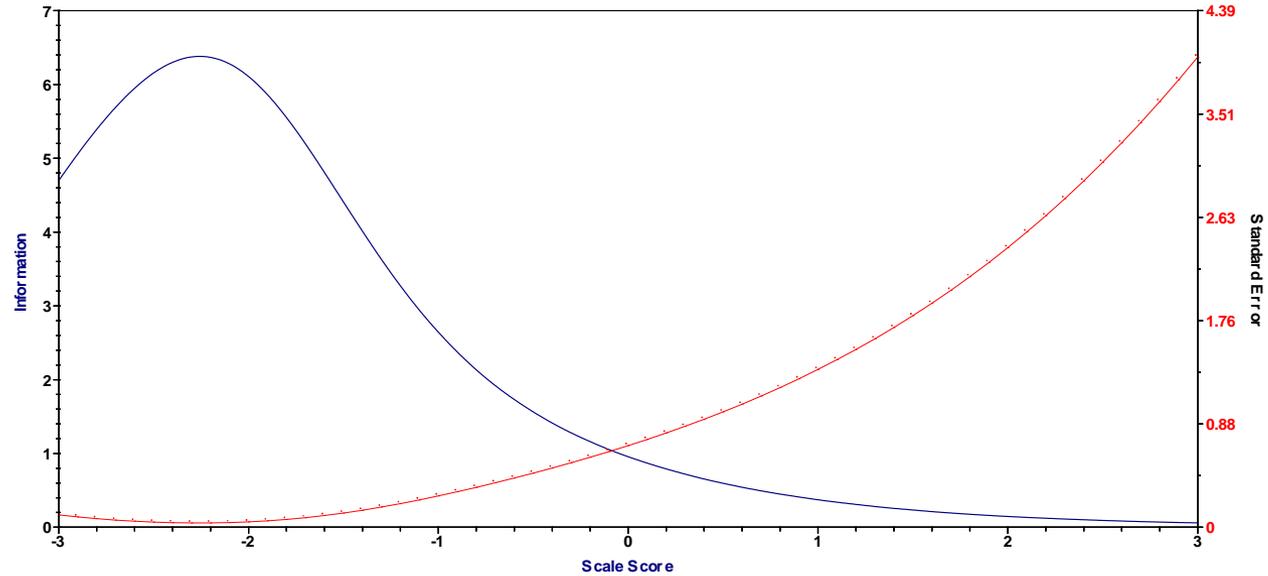
Abdominal pain

Abdominal Distension

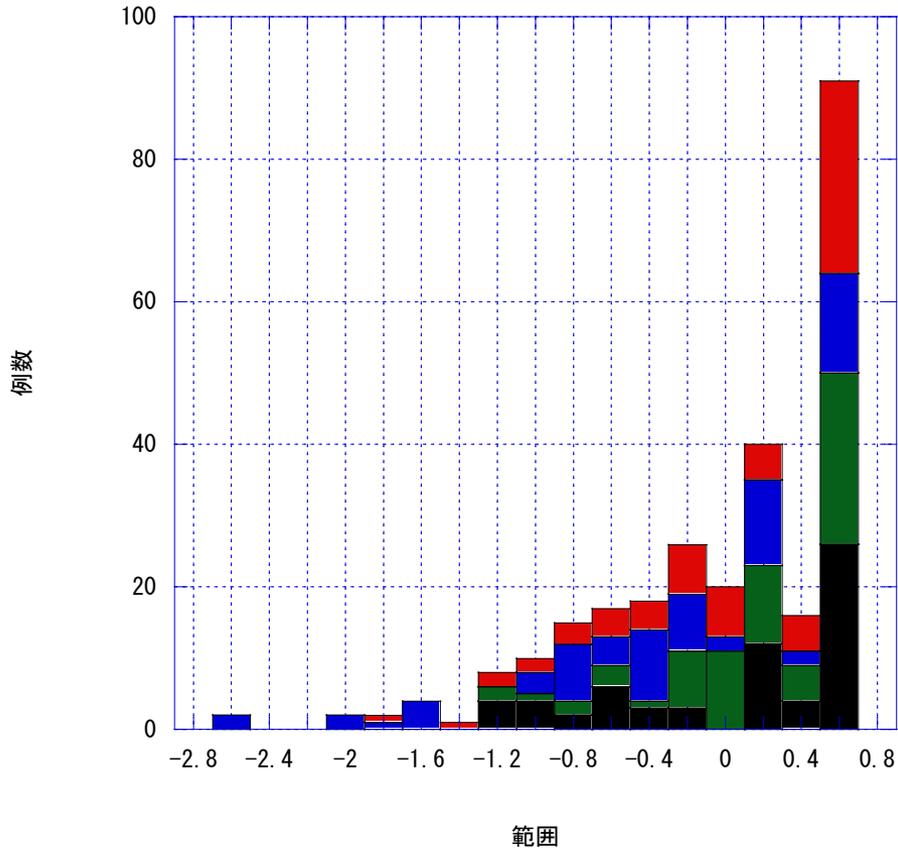
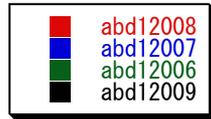


Abdominal distension

Abdominal Pain



The Difference of Examinee's Abilities for abdominal pain in 2006,7,8,9



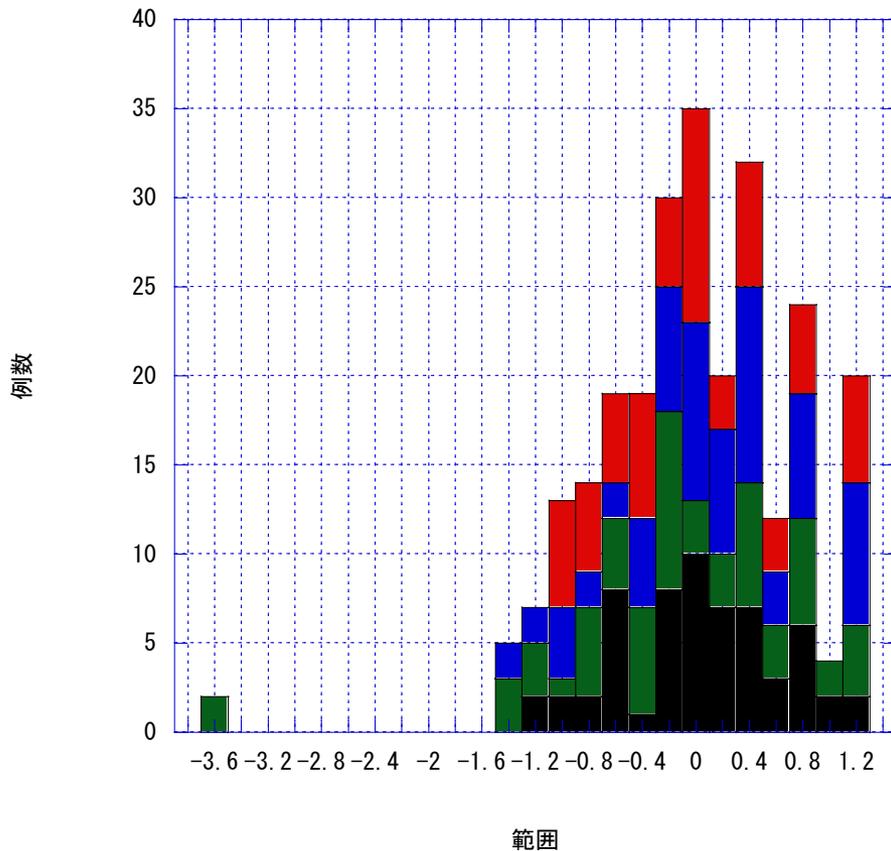
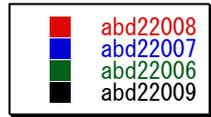
Sum of Squares 0.0492 128.4595
 Deg. of Freedom 1 270

Residual standard error: 0.689765
 Df Sum of Sq Mean Sq F Value
 Pr(F)

1 0.0492 0.0492456 0.1035058
 0.7479114

Residuals 270 128.4595 0.4757758

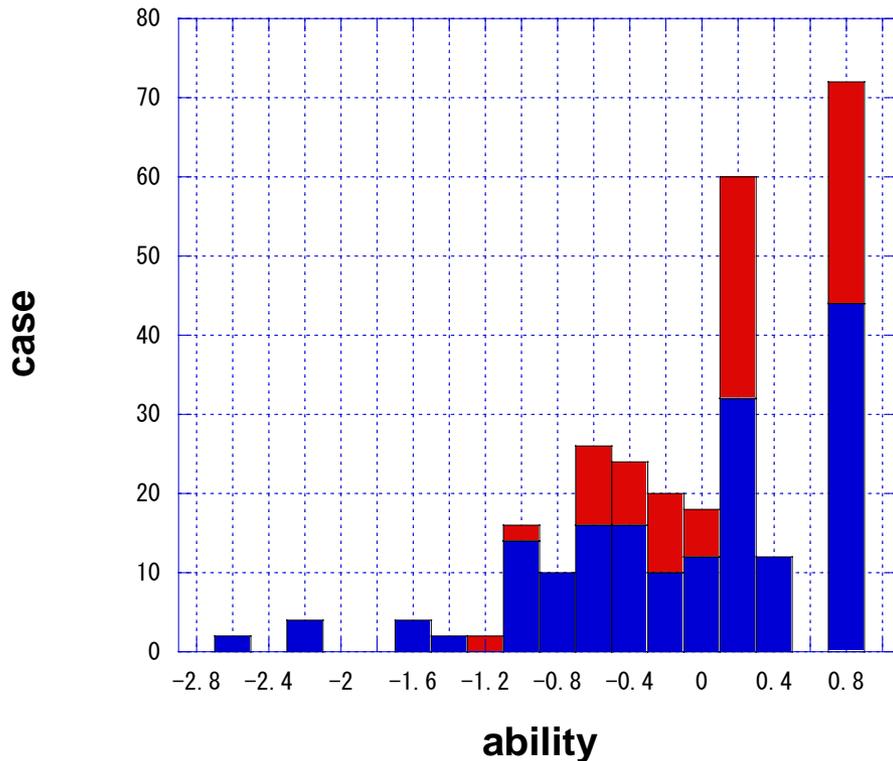
Difference of Examinee's Abilities for abdominal distension in 2006,7,8,9



Sum of Squares 0.5419 136.0932
Deg. of Freedom 1 254

Residual standard error: 0.7319836
Df Sum of Sq Mean Sq F Value Pr(F)
1 0.5419 0.5419465 1.011472 0.3155077
Residuals 254 136.0932 0.5358000

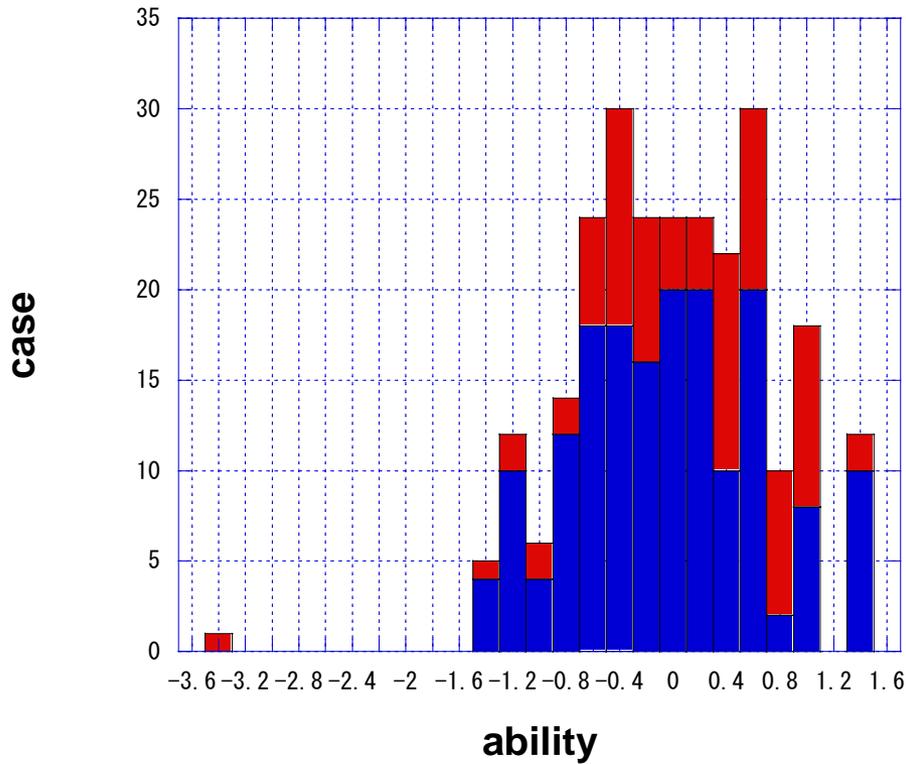
Difference of Examinee's abilities in gender for abdominal pain



Welch Modified Two-Sample t-Test
t = 2.882, df = 249.313171800256,
p-value = 0.0043

Blue male
Red female

Difference of Examinee's abilities in gender for abdominal distension



Welch Modified Two-Sample t-Test
 $t = 1.6376$, $df = 142.691144027503$,
 $p\text{-value} = 0.1037$

Blue male
Red female

Cause of the difference for abdominal pain

| ID | ITEM | correct-answer rate | | CONTRAST | SDIF | PROBABILITY | CHI-square |
|-----|---|---------------------|-------|----------|--------|-------------|------------|
| | | female | male | | | | |
| 6) | Appropriate eye contact | 0.968 | 0.944 | -0.026 | -0.051 | 0.411 | 0.677 |
| | | | | 0.511 | | | |
| 7) | Appropriate posture and attitude | 1 | 0.966 | 3.867 | 2.638 | 0.052 | 3.770 |
| | | | | 1.466 | | | |
| 14) | Ask for the level of the symptom | 0.628 | 0.697 | 0.788 | 0.589 | 0.221 | 1.495 |
| | | | | 1.338 | | | |
| 24) | Ask for the social situation | 0.936 | 0.815 | -0.815 | -2.322 | 0.064 | 3.436 |
| | | | | 0.351 | | | |
| 6) | Was the abdomen exposed enough? | 0.894 | 0.961 | 1.728 | 2.814 | 0.047 | 3.956 |
| | | | | 0.614 | | | |
| 12) | Inspection of important symptoms and expression of the findings | 0.947 | 0.944 | 0.512 | 0.813 | 0.184 | 1.768 |
| | | | | 0.63 | | | |
| 9) | Accurate percussion | 0.979 | 0.955 | -0.194 | -0.161 | 0.344 | 0.895 |
| | | | | 1.203 | | | |
| 15) | Proper superficial palpitation on the whole abdomen first | 0.989 | 0.933 | -1.375 | -2.795 | 0.047 | 3.935 |
| | | | | 0.492 | | | |
| 16) | Was the deep palpation performed correctly? | 0.947 | 0.921 | 0.121 | 0.299 | 0.292 | 1.109 |
| | | | | 0.405 | | | |
| 18) | Was the pain pressure point confirmed? | 0.862 | 0.697 | -0.671 | -1.912 | 0.083 | 2.998 |
| | | | | 0.351 | | | |
| 19) | Was rebound tenderness confirmed? | 0.83 | 0.753 | -0.068 | -0.300 | 0.292 | 1.110 |
| | | | | 0.227 | | | |

CONTRAST OF ITEM LOCATION

$$DIF(b2_i) = b2_i - b1_i$$

STANDARDIZED DIF

$$SDIF(b2_i) = \frac{DIF(b2_i)}{\sqrt{Var(b2_i) + Var(b1_i)}}$$

The IRT-available items were mostly different between “abdominal pain” and “abdominal distension”, and the IRT-available items were specific to the scenarios.

The values of item-difficulty and item-discrimination, and the examinee-ability values in OSCE, can be common-scaled, and equated horizontally and vertically over the past several years.