## يتوقع بعد الانتهاء من هذا العرض أن :

- نتعرف على مناحي القياس
- نتعرف على أنواع مؤشرات مطابقة الشخص .
- نتعرف برنامج WBstare
- نتعرف على خطوات تحليل بيانات (TIMSS ).
- توصيات لتحسين النتائج على اختبار (TIMSS)

## مناحي القياس

- ينتج عن قياس أي سمة إنسانية استجابات تترجم إلى أرقام وهو ما يعرف بعلم القياس النفسي هناك منحيان في علم القياس النفسي لدراسة تلك الاستجابات :
- الاول وفقاً للمجموعة التي خضعت لعملية القياس وهي النظرية التقليدية في القياس (Classical Testing Theory ) .
- الثاني أن تدرس تلك الاستجابات وفقاً لنموذج رياضي معين وهي نظرية استجابة الفقرة ( Item Response Theory ) ، وعليه فان لكل نظرية مؤشراتها الاحصائية الخاصة بها ، هذا ينسحب على مؤشرات المطابقة سواء للشخص أو الفقرة .

كيف يمكن النظر الى القياس النفسي (الاختبار)؟

يمكن النظر الى القياس النفسي (الاختبار ) : عبارة عن صراع بين قدرة المفحوص وصعوبة الفقرة ، في حال تغلبت القدرة على الصعوبة النتيجة (١) للمفحوص(٠) للفقرة، وإذا تغلبت صعوبة الفقرة على قدرة المفحوص فان النتيجة (١) للفقرة (٠) للمفحوص .

## مؤشرات الاحصائية لمطابقة الشخص :

- زاد الاهتمام في السنوات الاخيرة من أخصائي القياس النفسي بمؤشرات مطابقة الشخص ، وهناك تسميات عديدة في الادب النفسي لهذه المؤشرات مثل إحصائيات مطابقة الشخص (person fit statistics) مؤشرات الملائمة (appropriateness indices) مؤشرات التحذير ( Caution indices) .
- وبلغ عدد مؤشرات مطابقة الشخص ٣٦ مؤشرا وفق دراسة كاراباتسوس (Karabatsos) وتنقسم إلى مؤشرات المعلمية وللامعلمية .

# أسباب حدوث انماط الاستجابة الشاذة

- حصر ( Iasaonas,Bill&David,2002) أسباب حدوث الانماط الشاذة :
- الخصائص الشخصية (التخمين ، الغش ، القلق ، الدافعية أو عدم الاهتمام ، التباطؤ في أداء الاختبار ، التكاسل )
- الخصائص التربوية ( سوء فهم الفقرة ، الابداع الاستثنائي ،الاختلاف في أساليب التدريس ).
- عوامل خارجية ( الاخطاء في تعبئة ورقة الاجابة ، المرض )
- التحيز الثقافي ، التحيز العِرقي ، التحيز للجنس .

ملاحظة : يبقى تحديد السبب الحقيقي من الامور الصعبة لتدخل عدة عوامل في توثر اجابة المفحوص عن فقرات الاختبار

# المؤشرات اللامعلمية لمطابقة الشخص

- ويعرف المؤشر مطابقة الشخص (اللامعلمي) بأنه مؤشر إحصائي يحدد البعد أو المدى بين البيانات الفعلية _ استجابات المفحوصين _ القيم المتوقعة بناءً على المجموعة التي ينتمي اليها المفحوص ، للحكم على الانماط الاستجابة الشاذة غير المطابقة تقارن قيم مؤشر المطابقة مع نمط الاجابة المثالي لجتمان (Guttman) .

# المؤشرات اللامعلمية

## عائلة مؤشرات ضمن القدرة خارج نطاق القدرة ( BW)

- معظم المؤشرات الاحصائية المبنية على المجموعة معتمدة على نموذج جتمان (Guttman) لذلك إي نمط إجابة لا يتماشى من نموذج جتمان يعتبر شاذاً .

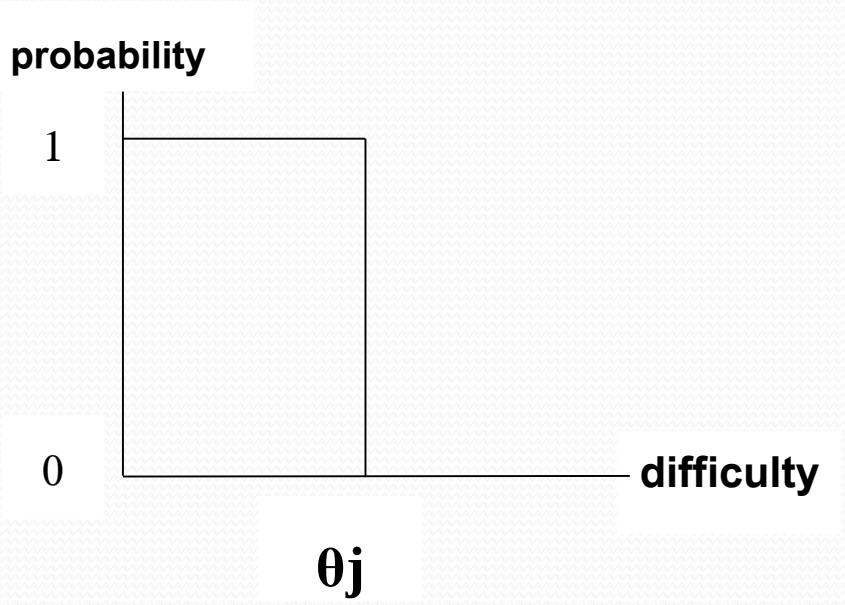
- طوّر هونغ ( Huang,2011,2012 ) مؤشرات للكشف عن الانماط الاستجابة مشتقة من مؤشر دكوستا ( D'Costa,1993a,1993b ) الذي أشتق بالأصل من مؤشر ساتو ( Cij ).

$$C_{ij} = 1 - \frac{\mathrm{cov}[u_{ij}, q_{.j}]}{\mathrm{cov}[u_j^{\$}, q_{.j}]}$$

**حيث أن (i) ترمز إلى الشخص ، ( j) يرمز إلى الفقرة .**

- **$u_{ij}$ : نمط الاجابة الملاحظ للمفحوص (٠ ، ١ ) .**
- **q.j : عدد الاصفار للفقرات الى عدد الاشخاص مرتبة من الادنى إلى الاعلى .**
- 
- **$u_i^{\$}$ : متجه جتمان العكسي للمفحوص ( i)**

## دالة استجابة الشخص ( Person Response Function) أو دالة الخطوة (Step Function ) أو مقياس جتمان




أن القدرة ثابتة والصعوبة متغيرة على متصل الصعوبة ، إذا كانت صعوبة الفقرات أكبر قدرة الشخص ( θj ) أو العلامة الحقيقية فأن إجابة الشخص تساوي (0) ، وإذا كانت صعوبة الفقرات أقل من ( θj ) فأن إجابة الشخص تساوي واحد (1) وبهذا تحصل على ما يعرف بالنمط الاجابة المثالي .

# نمط الاجابة المثالي

| م | ١ | ٢ | ٣ | ٤ | ٥ | المجموع |
|---|---|---|---|---|---|---|
| ١ | ١ | ١ | ١ | ١ | ١ | ٥ |
| ٢ | ١ | ١ | ١ | ١ | ١ | ٥ |
| ٣ | ١ | ١ | ١ | ١ | ٠ | ٤ |
| ٤ | ١ | ١ | ١ | ١ | ٠ | ٤ |
| ٥ | ١ | ١ | ١ | ٠ | ٠ | ٣ |
| ٦ | ١ | ١ | ١ | ٠ | ٠ | ٣ |
| ٧ | ١ | ١ | ٠ | ٠ | ٠ | ٢ |
| ٨ | ١ | ١ | ٠ | ٠ | ٠ | ٢ |
| ٩ | ١ | ٠ | ٠ | ٠ | ٠ | ١ |
| ١٠ | ١ | ٠ | ٠ | ٠ | ٠ | ١ |
| ١١ | ٠ | ٠ | ٠ | ٠ | ٠ | ٠ |
| المجموع | ١٠ | ٨ | ٦ | ٤ | ٢ | ٣٠ |

## مثال : افترضي يوضح بعض الانماط الاستجابة الشاذة

| | صعوبة الفقرات | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| السلوك | ١٠ | ٩ | ٨ | ٧ | ٦ | ٥ | ٤ | ٣ | ٢ | ١ | الشخص |
| تخمين | ١ | ١ | ٠ | ٠ | ٠ | ٠ | ١ | ١ | ١ | ١ | ١ |
| ابداع استثنائي | ١ | ١ | ١ | ١ | ١ | ١ | ٠ | ٠ | ٠ | ٠ | ٢ |
| غش | ١ | ٠ | ١ | ٠ | ١ | ٠ | ٠ | ٠ | ١ | ١ | ٣ |
| | 0.10 | 0.15 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | |

# مؤشر التمكّن ( capability index ).

- **مؤشر التمكّن ( capability index ). يرمز له $W_i^1$**
- **سمي بذلك لأنه يقيس الاستجابات الصحيحة عن الفقرات التي تكون مستوى صعوبتها أقل من قدرة المفحوص ( $i$ ). بمعنى أن المفحوص قد تمكّن من الفقرات التي تكون أقل من قدرته .**
- يعطى وفق الصيغة التالية

$$W_i^1 = \frac{\sum_{j=1}^{Ti} (u_{ij}) \times (q_{iT}^{*} - q_{.j})}{[(k-1)/2]} \times 100$$

# مؤشر سوء الفهم ( Misconception Index )

- مؤشر سوء الفهم ( Misconception Index ) يرمز له ( $B_i^0$ )
- سمي بذلك الاسم لأنه يقيس الاستجابات الخاطئة عن الفقرات التي تكون مستوى صعوبتها أعلى من قدرة المفحوص ( $i$ ) ، بمعنى أنه يقيس النقاط العمياء للطالب التي خارج قدرته التي فعلاً يكون هناك سوء فهم لدى الطالب .
- ويعطى وفق المعادلة التالية

$$B_i^0 = \frac{\sum_{j=T_i+1}^{K} (1-u_{ij}) \times (q_{.j} - q_{iT}^*)}{[(k-1)/2]} \times 100$$

# مؤشر اللامبالاة ( carelessness Index )

- مؤشر اللامبالاة ( carelessness Index ) يرمز له ( $W_i^0$ )
- سمي بذلك الاسم لأنه يقيس الاستجابات الخاطئة عن الفقرات التي تكون مستوى صعوبتها أقل من قدرة المفحوص ( $i$ ) .
- ويعطى وفق الصيغة التالية :

$$W_i^{\ 0} = \frac{\sum_{j=1}^{Ti} (1-u_{ij}) \times (q^*_{iT} - q_{.j})}{[(k-1)/2]} \times 100$$

# مؤشر التخمين ( Guessing Index )

- مؤشر التخمين ( Guessing Index ) يرمز له ( $B_i^1$ )

- سمي بذلك الاسم لأنه يقيس الاستجابات الصحيحة عن الفقرات التي تكون مستوى صعوبتها أعلى من قدرة المفحوص ( $i$ ) .
- يعطى وفق الصيغة التالية

$$B_i^1 = \frac{\sum_{j=T_i+1}^{k} (u_{ij}) \times (q_{.j} - q_{iT}^{*})}{[(k-1)/2]} \times 100$$

حيث أن (i) ترمز إلى الشخص ، (j ) يرمز إلى الفقرة .

- **uij** : نمط الاجابة الملاحظ للمفحوص (٠ ، ١ ) .
- **q.j** : عدد الاصفار للفقرات الى عدد الاشخاص مرتبة من الادنى إلى الاعلى .
- $q^*_{iT}$ : معدل النسب الاخضاع للفقرة ( Tth) و للفقرة (T+1)th للمفحوص الذي حصل على مجموع (T)
- لعدد فقرات K

# البرنامج (WBstar)

- كُتب برنامج وبستار(WBStar) من قبل لو و هوانغ (Lu&Huang,2010) باستخدام برنامج فيجوال بيسك ٦ (Visual Basic 6.0).كما أشار له هوانغ و يورو ( Huang&Wu ,2012):
- كيف يعمل البرنامج :
- ١. يتم ترتيب المفحوصين ترتيبا تنازلياً وفق الرتب المئنية ( percentiles) للمؤشر ( C) لأنه يعطي القدرة الحقيقية خالية من (التخمين ، اللامبالاة ) إلى ثلاثة أصناف (H,M,L ) ( H~ 0.67, M~(0.33-0.67),L~0.33) .
- ٢. تحسب مؤشرات مطابقة الشخص (W,B,C,M)
- ( $W_i^1$:$B_i^0$ ):(C,M)=1~Nomal.
- ($W_i^0$ ) :(W)=2~Carlessness
- $B_i^1$ ( B )=3~Geuessing
- W+B= 4 )~(Complex
- ٣. مؤشرات النظرية التقليدية في القياس .

# المؤشرات المعلمية  مطابقة الشخص

- ويعرف المؤشر مطابقة الشخص (المعلمية)  بأنه مؤشر إحصائي يحدد البعد أو المدى بين البيانات الفعلية _ استجابات المفحوصين _ القيم المتوقعة من خلال نموذج المستخدم الرياضي المستخدم ( IRT ) ، للحكم على الانماط الاستجابة الشاذة غير المطابقة  تقارن قيم مؤشر المطابقة مع قيمة حرجة (±1.96) .

## المؤشرات الاحصائية المبنية على البواقي(Wright,Master,1982) : مؤشر المطابقة الخارجي (Outfit) مؤشر المطابقة الداخلي ( infit) .

**يعني المؤشر المطابقة الداخلية : يتصف هذا المؤشر بحساسيته في الكشف عن الاجابات غير المتوقعة لمستوى صعوبة الفقرات القريبة من مستوى قدرة المفحوص .**

$$ZW = [(Ms^{\frac{1}{3}} - 1)] \left(\frac{3}{\sigma}\right)] + [(\frac{\sigma}{3})$$

$$MS = \frac{\sum_{i=1}^{k}[X_i - P_i(\theta)]^2}{\sum_{i}^{k} w_{ij}}$$

$$\sigma = \frac{\sum_{i=1}^{k} w_{ij} - 4\sum_{i}^{k} w_{ij}\ ]^{1/2}}{\sum_{i=1}^{k} w_{ij}}$$

- **Xij: متجه الاجابات الملاحظة للمفحوص(j ) على مجموعة من الفقرات (i)**
- **Pij($\theta$ ): متجه احتمالات الاجابات للمفحوص(j ) على مجموعة من الفقرات (i) بشكل صحيح .**
- **Wij: تشير تباين الفقرة (i) وتساوي [Pi($\theta$ )][1-Pi($\theta$ )]**

## المؤشرات الاحصائية المبنية على الارجحية العظمي (Hulin,Drasgow&Parson,1983) : Lz

- تعتمد قيم المؤشر على قيم معالم القدرة والفقرات المقدرة والتي تمثل أنماط الاستجابة الملاحظة بأعلى احتمال يمكن الوصول اليه ضمن نموذج من نماذج الاستجابة للفقرة ، وعليه إذا لم تكن هناك قيمة للقدرة المقدرة تعظم دالة الارجحية العظمى لاستجابات مفحوص ما عندها تكون قيمة القدرة المقدرة غير ملائمة لهذا المفحوص وبالتالي لا تمثل إجابته .

$$l_z = \frac{l_o - E(l_o)}{[\mathrm{var}(l_o)]^{1/2}}$$

$$l_o = \sum_{i=1}^{k} \{X_i \ln p_i(\theta) + (1 - X_i) \ln[1 - p_i(\theta]\}$$

$$E(l_o) = \sum_{i=1}^{k} \{p_i(\theta)[\ln p_i(\theta)] + [1 - p_i(\theta)] \ln[1 - p_i(\theta)]$$

$$V(l_o) = \sum_{i=1}^{k} \{p_i(\theta)[\ln p_i(\theta)] \left[ \ln \frac{p_i(\theta)}{1 - p_i(\theta)} \right]$$

# المؤشرات الاحصائية المبنية على التباين المشترك

- **طوّر لين وتاتسوكا (Linn&Tastouka,1983) مؤشرات مطابقة مبنية على التباين المشترك وكانت امتدادا لمؤشر ساتو ( Sato) . حيث استخدم النموذج الاحتمالي في بناء مصفوفة الاحتمالات ومصفوفة العلامات الخام التي رتبت فيهما الفقرات من الاسهل الى الأصعب والمفحوصين من القدرة العالية الى القدرة المنخفضة . وقد أشتق ستة مؤشرات غير معيارية مبنية على التباين المشترك .**

$$\xi_2 = \frac{\sum_{j=1}^{k}(P_{ij} - X_{ij})(P_{ij} - T_i)}{[\sum_{j=1}^{k} P_{ij}Q_{ij}(P_{ij} - T_i)^2]^{1/2}}$$

- $X_{ij}$: متجه الاجابات الملاحظة للمفحوص( j ) على مجموعة من الفقرات (i)
- $P_{ij}$: متجه احتمالات الاجابات للمفحوص( j ) على مجموعة من الفقرات (i) بشكل صحيح .
- $Q_{ij}$: متجه احتمالات للمفحوص( j ) على مجموعة من الفقرات (i) بشكل خاطئ
- $T_i$: متوسط متجه الاحتمالات في المصفوفة بشكل أفقي .

# دراسات لمقارنة بين المؤشرات المعلمية واللامعلمية

- أظهرت الدراسات السابقة تفوق مؤشرات المطابقة المبنية على المجموعة في قوتها في الكشف عن الانماط الشاذة عن المؤشرات المبنية على نموذج ( IRT) في دراسة ( Karabatsos,2003) الذي قارن فيها بين ٣٦ مؤشر اً .
- وقد أجرى هونغ عدة دراسات عن عائلة مؤشرات WB حيث أجرى دراسة للمقارنة بين مؤشرات المبنية على المجموعة ( SCI,MCI,NCI,WB WB,) مع مؤشرات المبنية على نموذج IRT ( OUTFITz , INFITz, ECI2z,ECI4z,Lz ) تحت ظروف مختلفة أظهرت النتائج أن المؤشرات (WB ) قد كشف أكثر 90% من الانماط الشاذة وكانت أكثر استقراراً تحت الظروف المختلفة .

# توظيف القياس في تشخيص المعرفي

# Cognitive Diagnostic assessment

- يعتبر توظيف القياس تشخيص المعرفي ( CDA ) من التوجهات الحديثة في البحث أول من استخدم ( Tatsoka,1984) في تشخيص لأخطاء في المهارات الرياضية الفرعية وذلك من خلال ( Rule Space ).
- كتاب Gierl,Leighton(2007) "Cognitive Diagnostic assessment for education "

- حرر كتاب من قبل دراسكوا (٢٠١٥ ) ( Technology and Testing: Improving Educational and Psychological Measurement ) يتضمن فصل :
- "Diagnostic Assessment Methods for the Reliable Measurement of Multidimensional Abilities . ( DAM) .
- استخدم هونغ (Haung,2012 ) مؤشرات (WB) في تشخيص الأخطاء لاختبار الرياضيات الذي يقوم بأعداده المعلم
- " Classroom-based Cognitive Diagnostic Model For a Teacher-made Fraction-Decimal Test"
- صمم برنامج ( BULIB,1989) من قبل تاتسوكا لتشخيص الاخطاء في المقاييس الكبيرة مثل SAT و TOEFL باستخدام (Rule Space) .

# اختبار (TIMSS)

- يعقد الاختبار الدولي للعلوم والرياضيات ( TIMSS) ( Trends In International Mathematics and Science) منذ عام ١٩٩٥ بشكل دوري كل أربع سنوات ، ويهدف إلى مساعدة دول العالم في تحسين تعلم الطلبة في العلوم والرياضيات من خلال تزويدها بالتغذية الراجعة على أداء الطلبة واستطلاع أراء المعلمين ومدراء المدارس .
- وتم استخدام بيانات (Timss,2015) للنموذج الثاني لمبحث العلوم ، وقد اشترك في الاختبار ٥١ دولة .

# خطوات تحليل البيانات

١. تنزيل البيانات عن الموقع الالكتروني
٢. https://timssandpirls.bc.edu/
٣. ٢. اختيار الملفات إجابات الطلاب التي تحمل رمز ( BAS.... ) .
٤. ٣.تجميع الملفات في ملف واحد ( Materfile)
٥. ٤. فرز البيانات حسب رقم النموذج .
٦. ٥.تصحيح كل نموذج على حدى من خلال ملف التصحيح BSASCRM6
٧. ٦. فرز الملفات حسب الدول المشاركة
٨. ٧. استبعاد الفقرات متعددة التدريجات .
٩. ٨. تطبيق مؤشر الاحصائي باستخدام برنامج WBstare

# نماذج_فقرات_TIMSS

| النموذج | عدد فقرات الرياضيات | عدد_الفقرات_المتعددة | عدد فقرات العلوم | عدد_فقرات_المتعددة |
|---|---|---|---|---|
| النموذج_١ | ٣٥ | ١٩ | ٣٩ | ٢٣ |
| النموذج_٢ | ٣٢ | 20 | 47 | 28 |
| النموذج_٣ | 28 | 15 | ٤٣ | 25 |
| النموذج_٤ | ٣٢ | ١٣ | ٣٧ | ١٩ |
| النموذج_٥ | ٣٤ | 15 | 32 | ١٧ |
| النموذج_٦ | 32 | ١٧ | ٣٤ | ١٩ |
| النموذج_٧ | ٣٢ | ١٧ | 32 | ١٧ |
| النموذج_٨ | 30 | 16 | ٢٨ | 14 |
| النموذج_٩ | ٢٩ | ١٦ | ٣٤ | ٢٠ |
| النموذج_١٠ | ٢٩ | 17 | ٤٠ | 26 |
| النموذج_١١ | 29 | ١٦ | ٤٢ | ٢٧ |
| النموذج_١٢ | ٣٠ | ١٨ | ٤١ | ٢٤ |
| النموذج_١٣ | ٣٦ | 21 | ٤٠ | ٢٢ |
| النموذج_١٤ | ٣٠ | 16 | ٣٩ | ٢٠ |

# مراجعة الملفات

- مراجعة ملف الاكسل :
- فقرات الاختبار
- التحليل

# جدول المواصفات لاختبار العلوم

| بعد المحتوى | النسبة | البعد المعرفي | النسبة |
|---|---|---|---|
| الحياة | %٣٥ | المعرفة | %٣٥ |
| الكيمياء | %٢٠ | التطبيق | %٣٥ |
| الفيزياء | %٢٥ | الاستنتاج | %٣٠ |
| علوم الارض | %٢٠ | | |

| Class | سنغافورة | سعودية |
|---|---|---|
| H1 | ١٤٧ | ١٨ |
| H1' | ٢ | ٢ |
| H3 | ١٧ | ٠ |
| M1 | ٣١ | ٢٢ |
| M1' | ٣٩ | ٦٧ |
| M2 | ٤ | ٦ |
| M2' | ١ | ١ |
| M3 | ٠ | ١ |
| M4 | ٠ | ١ |
| L1 | ٢١ | ١٣٧ |
| L2 | ٦ | ١١ |
| L2' | ١ | ١ |
| L3 | ٠ | ٢ |
| L3' | ٠ | ١ |
| L4 | ١ | ٠ |
| Normal | ٢٤٠ | ٢٤٦ |
| Guessing | ١٢ | ١٩ |
| Careless | ١٧ | ٤ |
| Complex | ١ | ١ |
| Total | ٢٧٠ | ٢٧٠ |

# الاستنتاجات بناء على التحليل

- يلاحظ أن النسبة الكبرى من أنماط الاجابات (٩١%) من الطلبة السعوديين و (٨٨%) من الطلبة السينغافوريين حيث صنفت كأنماط استجابة طبيعية هذا يعني أن الدرجات التي حصلوا عليها الطلاب تمثل مستواهم الحقيقي .
- يلاحظ أن (٧%) من إجابات الطلبة السعوديين صنفت كأنماط استجابة أعتمد فيها الطلبة على التخمين (٤%) من إجابات الطلبة سينغافوريين صنفت كأنماط أعتمد فيها الطلبة على التخمين .
- يلاحظ أن (١%) من إجابات الطلبة السعوديين صنفت كأنماط استجابة غير مبالية ، (٦ %) من إجابات الطلبة سينغافوريين صنفت كأنماط استجابة غير مبالية .
- نمط من الاجابة (١) صنف كنمط مقعد ( جمع بين التخمين و اللامبالاه ) للطلبة السعوديين ونمط (١) للطلبة سينغافورين .

# كيف يمكن زيادة تحصيل الطلبة

- https://www.youtube.com/watch?v=YaXYLvBmA6U
- تعزيز الدافعية ومفهوم الذات ( Self-Concept, self-confidence) لدى الطلبة والمعلمين لما لها من دور كبير في زيادة التحصيل حيث هناك كثير الدراسات التي أيدت ذلك .
- A comparative investigation of the presence of psychological conditions in high achieving eighth graders from TIMSS 2007 Mathematics,2012 by
- Analyzing Affective Factors Related to Eighth Grade Learners' Science and Mathematics Achievement in TIMSS 2007,2013
- Multilevel Effects of Teacher Characteristics on TIMSS 2011 Science Achievement,2011

- توظيف التقنية الحديثة بفعالية بحيث يتم الاستفادة من مواقع التشاركية WiKi, Facebook ... ( طرح الاسئلة والاجابات من خلال هذه المواقع ) وفق نموذج (Model Collaborative learning ) .
- **Effects of social learning networks on student academic achievement and pro-social behavior in accounting,2017**

- أظهرت نتائج الدراسة أن الحالة الاقتصادية والثقافية (Socio-Ecomonic) للعائلة والمدرسة كان لها دور كبير في تحصيل الطلبة ، كذلك أن الاداء في المدارس الخاصة كان أفضل .

- **Context factors and student achievement in the IEA studies: evidence from TIMSS**
- **THE RELATIONSHIP OF STUDENT DISPOSITIONS AND TEACHER CHARACTERISTICS WITH THE MATHEMATICS ACHIEVEMEN STUDENTS IN LEBANON AND SIX ARAB COUNTRIES IN TIMSS 200**