

IT 201

Chapter 04

Evaluating interface Designs

مقدمة

- المصممين قد يصبحوا جدا مبهجين (so entranced) مع ابداعهم (creations) التي قد تفشل تقييمهم بشكل كاف (adequately).
- المصممين المحنكين (Experienced designers) بلغوا الحكمة والتواضع (have attained the wisdom and humility) ليعرفوا أن الاختبارات الممتدة أمر ضروري.
- المحددات لخطة التقييم (The determinations of the evaluation plan) يتضمن:
 - 1- مرحلة التصميم (stage of design): في وقت مبكر (early)، منتصف (middle)، في وقت متأخر (late).
 - 2- حداثة المشروع (novelty of project): معرفة جيدا مقابل استكشافي (well defined vs. exploratory).
 - 3- عدد المستخدمين المتوقعين.
 - 4- حرجية الواجهة (criticality of the interface): نظام طبي للحياة الحرجة (life-critical medical) مقابل دعم معرض المتحف (vs. museum exhibit support).
 - 5- تكاليف المنتج (costs of product) والتخصيص المالي (finances allocated) للاختبار.
 - 6- الوقت المتاح.
 - 7- تجربة التصميم وتقييم الفريق.
- المقيمين المستخدمين (Usability evaluators) يجب عليهم توسيع اساليبهم (broaden their methods) ويكون منفتح للأساليب الغير تجريبية، مثل رسومات المستخدم (user sketches)، واعتبار بدائل التصميم (consideration of design alternatives)، ودراسات الاثنوغرافيك (ethnographic studies).
- توصيات (Recommendations) تحتاج لأن تكون مبنية على النتائج الرصدية (observational findings).
- فريق التصميم يحتاج بأن يكون مشارك مع الأبحاث على عقبات (drawbacks) تصميم النظام الحالي.
- الأدوات والتقنيات تكون متطورة (are evolving).
- نطاق التكاليف () قد تكون 20% من المشروع وصولا إلى 5%.
- اختبار قابلية الاستخدام اصبح جزء مؤسس ومقبول (established and accepted) من عملية التصميم.

مراجعات الخبراء (Expert Reviews):

- في حين العروض الغير رسمية (While informal demos) للزملاء والعملاء يمكن توفير بعض ردود الفعل المفيدة، وقد أثبتت مراجعات الخبراء الغير رسميين بأنها فعالة.
- مراجعات الخبراء يستلزم (entail) نصف يوم من جهد اسبوع واحد، بالرغم من طول فترة التدريب ففي بعض الاحيان قد يتطلب شرح مجال المهام (task domain) أو الاجراءات التشغيلية.

- هناك أساليب متنوعة من مراجعة الخبراء للاختيار من:
 - 1- تقييم ارشادي (Heuristic evaluation).
 - 2- مراجعة الارشادات (Guidelines review).
 - 3- التفتيش المتناسق (Consistency inspection).
 - 4- التجول المعرفي (Cognitive walkthrough).
 - 5- استعارات من التفكير البشري (Metaphors of human thinking).
 - 6- تفتيش قابلية الاستخدام الرسمي (Formal usability inspection).
- مراجعات الخبراء بإمكانها أن تكون مجدولة عند عدة نقاط في عملية التطوير وذلك عندما يكونوا الخبراء متاحين وفريق التصميم جاهزين لردود الفعل.
- مختلف الخبراء يميلون إلى (tend) ايجاد اخطاء مختلفة في الواجهة، لذا 3-5 من مراجعات الخبراء تكون مثمرة للغاية (highly productive)، حيث يمكن اختبار قابلية الاستخدام المكمل (complementary).
- المخاطر مع مراجعات المستخدم هو أن الخبراء قد لا يكون لديهم الفهم الكافي (adequate understanding) لمجال المهام أو مجتمعات المستخدم.
- حتى مراجعات الخبراء المحنكين (Even experienced expert reviewers) لديهم صعوبة عظيمة في معرفة كيفية المستخدمين النموذجيين (typical users)، وخاصة المستخدمين لأول مرة سوف يتصرفون حقا.

ارشاد قابلية الاستخدام خطوة بخطوة (Step-by –step Usability Guide):

- A. الخطة (Plan).
- B. التحليل (Analyze).
- C. التصميم (Design).
- D. الاختبار والتصفية (Test and Refine).

اختبار قابلية الاستخدام والمختبرات (Usability Testing and Laboratories):

- بزوغ (The emergence) قابلية الاستخدام والمختبرات منذ أوائل الـ 1980s
- اختبار قابلية الاستخدام ليس فقط تسريع العديد من المشاريع ولكنه انتج حفظ تكاليف دراماتيكية.
- التحرك نحو (Movement towards) اختبار قابلية الاستخدام يحفز (stimulated) بناء مختبرات قابلة للاستخدام.
- مختبر قابلية الاستخدام البسيط النموذجي (A typical modest) يكون لديه مساحة 10 في 10 قدم، واحد من المنتسبين يقوم بعملهم والآخر مفصول بمرآة نصفها من الفضة للقائمين بالاختبار والمراقبين.
- المنتسبين ينبغي عليهم الاختيار لتمثيل مقاصد مجتمعات المستخدم، مع الانتباه لـ :
 - الخلفية في تجربة الحوسبة مع المهام، التحفيز (motivation)، والتعليم، والقدرة مع اللغة الطبيعية المستخدمة في الواجهة.
- المنتسب ينبغي عليه دائما متطوع (voluntary) والحصول على الموافقة المسبقة (and informed consent should be obtained).
- الممارسة الاحترافية هو أن يسأل على جميع مواضيع القراءة وتوقيع الافادة (statement) مثل:
 - a) أنا متطوع حرا (I have freely volunteered) للالتساب في هذه التجربة.
 - b) أنا على علم مسبق في ماذا سيكون مهامي وما هي الاجراءات المتبعة.
 - c) لقد اعطيت الفرصة لأسأل الاسئلة واجاباتها نالت على رضاي.

- (d) أنا على وعي ولدي الحق في سحب الموافقة (to withdraw consent) وعدم الاستمرار في المشاركة في أي وقت، بدون المساس في علاج مستقبلي (without prejudice to my future treatment).
- (e) توقيعي في الأسفل قد يؤخذ كتأكيد (as affirmation) لجميع البيانات المذكورة في الأعلى؛ وقد كانت معطاه من قبل للمشاركة في هذه الدراسة.
- اشربة الفيديو لمهام اداء المنتسبين تكون عادة قيمة للمراجعة في وقت لاحق وذلك لعرض المخاطر ومديري المشاكل التي تواجه المستخدمين:
 - استخدام التحذير (Use caution) لكي لا يتداخل (in order to interfere) مع المنتسبين.
 - دعوة المستخدمين بالتفكير بصوت عالي (aloud) "بعض الاحيان يشير إلى التفكير المتزامن مع الصوت العالي".
 - العديد من الأشكال المتنوعة (variant forms) لاختبار قابلية الاستخدام تمت محاولته:
 - (a) ورق الـ mockups.
 - (b) تخفيض اختبار قابلية الاستخدام.
 - (c) اختبار قابلية الاستخدام التنافسي (Competitive usability testing).
 - (d) اختبار قابلية الاستخدام الكوني (Universal usability testing).
 - (e) حفل الاختبار والمختبرات المتنقلة (Field test and portable labs).
 - (f) اختبار قابلية الاستخدام عن بعد (Remote usability testing).
 - (g) هل يمكنك كسر هذا الاختبار (Can you break this test).

أدوات المسح (Survey Instruments):

- استبيانات المستخدم المكتوبة (Written user surveys) هو رفيق مألوف و غير غالي ومقبول بشكل عام (are familiar, inexpensive and generally acceptable) لاختبارات قابلية الاستخدام ومراجعات الخبراء.
- المفاتيح للاستبيانات الناجحة:
 1. أهداف واضحة مسبقاً.
 2. تطوير العناصر المركز عليها التي تساعد على تحقيق الاهداف (attain goals).
- أهداف الاستبيان يمكنها أن تكون مرتبطة (tied) مع مكونات الكائنات واجراء نموذج الواجهة لتصميم الواجهة.
- بإمكان المستخدمين ان يسألوا عن انطباعاتهم الذاتية (subjective impressions) حول جوانب معينة (specific aspects) للواجهة مثل تمثيل:
 - (a) كائنات واجراءات مجال المهمة.
 - (b) تركيب جمل المدخلات (Syntax of inputs) وتصميم شاشات العرض (design of displays).
- أهداف أخرى تتمثل في التأكد:
 - 1- خلفية المستخدمين: العمر، الجنس، الأصول، التعليم، الدخل.
 - 2- التجربة مع الكمبيوترات: تطبيقات محددة أو رزم برمجية، طول الوقت، عمق المعرفة.
 - 3- مسؤوليات العمل: تأثير (influence) اتخاذ القرار، الادوار الادارية (managerial roles)، التحفيز (motivation).
 - 4- نمط شخصي (personality style): انطوائي مقابل منفتح (introvert vs. extrovert)، مخاطر مقابل تجنب المخاطر، التين المبكر مقابل المتأخر (early vs. extrovert)، المنهجية مقابل الانتهازية (systematic vs. opportunistic).
 - 5- اسباب عدم استخدام الواجهة: خدمات غير مؤهلة (inadequate services)، معقد جدا، بطيئة جدا.

- 6- الألفة مع الميزات (familiarity with features): الطباعة، وحدات المايكرو، الاختصارات، البرامج التعليمية (tutorials).
- 7- شعورهم بعد استخدام الواجهة: مشوش مقابل واضح (confused vs. clear)، محبط مقابل متحكم (bored vs. excited)، ممل مقابل متحمس (frustrated vs. in-control).

استبيانات (Surveys):

- استبيانات الأون لاين تتجنب تكاليف الطباعة واحتياجات الجهود الإضافية في التوزيع وتجميع النماذج الورقية.
- العديد من الناس يفضلون الاجابات على الاستبيان الموجز المعروضة على الشاشة بدلا من التعبئة واعادة النماذج المطبوعة.
- على الرغم من احتمالية وجود تحيز في العينة (bias in the sample).
- مثال الاستبيان هو استطلاع لرضى تفاعل المستخدم (QUIS: Questionnaire for User Interaction Satisfaction).

اختبار القبول (Acceptance Test):

- لمشاريع التنفيذ الكبيرة، العميل أو المدير عادة يحدد أهداف موضوعية وقابلة للقياس (sets objective and measurable goals) لأداء الأجهزة والبرمجيات.
- لو تم فشل منتج مكتمل لتلبية معايير القبول، يجب على النظام اعادة الصياغة حتى يثبت النجاح (must be reworked until success is demonstrated).
- بدلا من المعايير الغامضة والمضللة (the vague and misleading) "للاستعمال الحسن"، المعايير القابلة لواجهة المستخدم تستطيع أن تكون مؤسسة لما يلي:
 1. الوقت لتعلم وظائف معينة.
 2. السرعة في اداء المهمة.
 3. معدل الاخطاء للمستخدم.
 4. الاحتفاظ البشري للأوامر على مر الزمن (Human retention of commands over time).
 5. رضى المستخدمين الشخصية (Subjective user satisfaction).
- في النظام الكبير، قد يكون هناك 8 أو 10 اختبارات تنفذ على مكونات مختلفة للواجهة ومع مجتمعات المستخدمين المختلفة.
- عند نجاح اختبار القبول، قد يكون هناك فترة من الاختبار الميداني (field test) قبل التوزيع الوطني أو العالمي.

التقييم خلال الاستخدام النشط (Evaluation During Active Use):

- نجاح الاستخدام النشط يتطلب الانتباه المستمر (constant attention) من المدراء المخصصين (dedicated managers)، وموظفي خدمة المستخدم (user-services personnel)، وموظفي الصيانة (maintenance staff).
- الكمال لا يمكن تحقيق (Perfection is not attainable)، ولكن التحسينات النسبية (but percentage improvements) تكون ممكنة.

- مقابلات ونقاشات جماعية مركزة (Interviews and focus group discussions):
 - المقابلات مع المستخدمين فرديا يمكن أن تكون منتجة لأن المقابلة قد تصل إلى قضايا معينة من الاهتمام (can pursue specific issues of concern).
 - مناقشات المجموعة تكون ذات قيمة لتأكيد التعليقات الشاملة (the universality of comments).
- تسجيل البيانات المستمر لأداء المستخدم (Continuous user-performance data logging):
 - هندسة البرمجيات (Software architecture) يجب ان تجعلها سهلة على مديري النظام لجمع البيانات حول:
 - (a) أنماط استخدام النظام (The patterns of system usage).
 - (b) سرعة اداء المستخدم.
 - (c) معدل الاخطاء.
 - (d) الطلب المتكرر (Frequency) للمساعدة أون لاين.
 - الفائدة الرئيسية هو ارشاد مشرفي النظام (is guidance to system maintainers) في تحسين الاداء (in optimizing performance) وتقليل التكاليف لجميع المشاركين.
 - الاستشارات (consultant) عبر التليفون أو الأون لاين، الرسائل الالكترونية وصناديق الاقتراحات عبر الاون لاين:
 - (a) العديد من المستخدمين يشعرون بالاطمئنان (feel reassured) لو عرفوا بأن هناك مساعدة بشرية متاحة (a human assistance available).
 - (b) في بعض أنظمة الشبكة، المستشارين (Consultants) بإمكانهم مراقبة (monitor) كمبيوتر المستخدم ويرى نفس العروض الذي يراها المستخدم.
- صندوق الاقتراح الاون لاين أو تقرير مشكلة في الايميل (Online suggestion box or email trouble :reporting):
 - الرسائل الالكترونية للمشرفين (for maintainers) أو المصممين.
 - لبعض المستخدمين، كتابة رسالة يمكن ان ينظر إليها أنها تتطلب الكثير من الجهد.
- مجموعات المناقشة، الويكي والمجموعات الاخبارية (wiki's & newsgroups):
 - (a) إرساليات مسموح بها (Permit postings) للرسائل والاسئلة المفتوحة.
 - (b) البعض يكون مستقل، مثل America Online والياهو!
 - (c) قائمة موضوع (Topic list).
 - (d) رؤساء الجلسات في بعض الاحيان (Sometimes moderators).
 - (e) أنظمة اجتماعية.
 - (f) التعليقات والاقتراحات ينبغي أن تكون مشجعة.

السيطرة على التجارب الموجهة نفسيا (Controlled Psychologically-oriented Experiments):

- العملية العلمية والهندسية (Scientific and engineering process) تكون غالبا محفزة (is often stimulated) بواسطة التقنيات المحسنة للقياسات الدقيقة (for precise measurement).
- العملية السريعة (Rapid progress) في تصاميم الواجهات ستكون محفزة (simulated) كباحثين وممارسين (as researchers and practitioners) بتطوير مقاييس اداء البشرية المناسبة والتقنية.
- ملخص الاسلوب العلمي (Outline of the scientific method) كالتطبيق في تفاعل الكمبيوتر البشري قد يتضمن (comprise) هذه المهام:
 - (1) التعامل مع المشاكل العملية (practical problem) واعتبار اطار النظرية (consider the theoretical framework).
 - (2) توضيح فرضية الشفافية والقابلة للاختبار (State a lucid and testable hypothesis).

- (3) تعريف عدد قليل من المتغيرات المستقلة (independent variables) التي تتم معالجتها (to be manipulated).
- (4) اختيار المتغيرات المستقلة التي تتم قياسها.
- (5) التحكم في العوامل المتحيزة (Control of biasing factors) مثل: عينة غير تمثيلية من المواضيع (non-representative sample of subjects) أو اختيار المهمات، اجراءات الاختبار الغير متناسق (inconsistent testing procedures).
- (6) تطبيق اساليب الاحصائية لتحليل البيانات.
- (7) حل المشكلة العملية (resolve the practical problem)، تنقية النظرية (refine the theory)، واعطاء نصيحة للباحثين في المستقبل (give advice to future researchers).
- التجارب المحكمة بإمكانها ضبط نهاية (fine tuning) واجهة كمبيوتر البشرية للأنظمة المستخدمة النشطة.
 - الاداء يمكن مقارنته مع مجموعة التحكم.
 - المقاييس المستقلة يمكن أن تشمل أوقات الأداء، والرضى الشخصي للمستخدم، ومعدلات الأخطاء، واحتفاظ المستخدم عبر الزمن (user retention over time).

End of Chapter 04