CritiqueKit: A Mixed-Initiative, Real-Time Interface For Improving Feedback

C. Ailie Fraser, Tricia J. Ngoon, Ariel S. Weingarten, Mira Dontcheva, Scott Klemmer

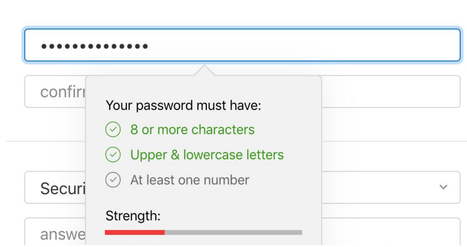
Giving Feedback is Hard

Feedback is a formative part of learning. However, good feedback is rare because it is time-consuming to give and people are not consistently skilled at providing it.

How can we guide reviewers toward improving the feedback they give?

What Inspired Us?

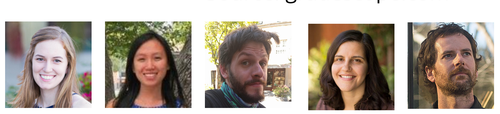
Password Validators



Source: appleid.apple.com

Feedback Reuse Systems

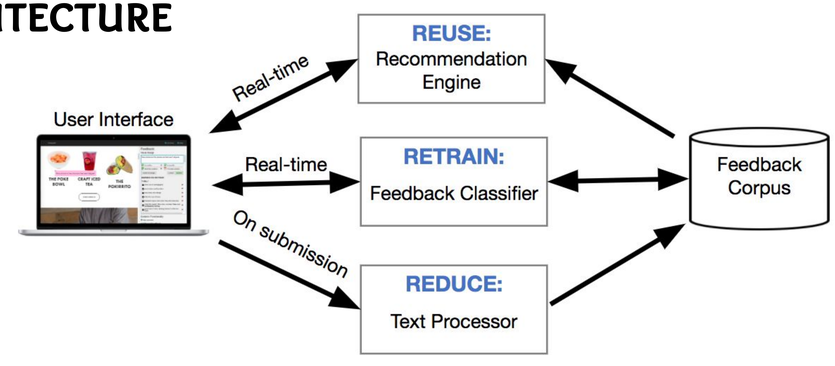
Source: gradescope.com



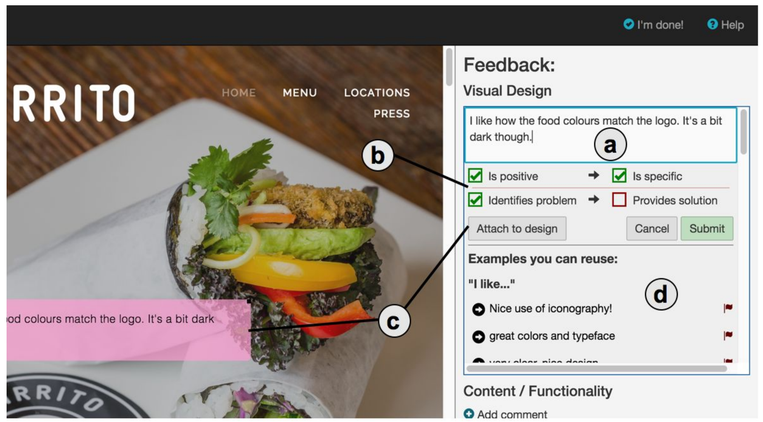
UC San Diego
The Design Lab

CritiqueKit Web Architecture

The CritiqueKit architecture comprises five components: a web interface, a corpus of feedback, a feedback classifier, a text processor, and a recommendation engine.



CritiqueKit introduces an approach for *reusing* prior feedback, *reducing* it to be useful in a general context, and *retraining* the system about what is useful in real time.



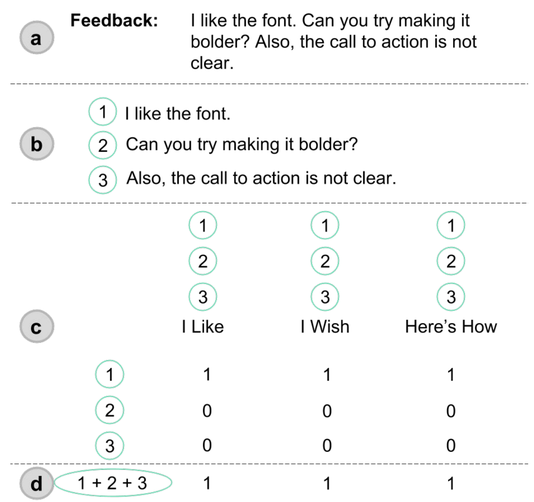
a) The reviewer types their feedback.

b) Checkboxes update in real-time showing which criteria of good feedback the comment fulfills.

c) The reviewer can attach their feedback to the document being reviewed.

d) The reviewer can browse and reuse suggested feedback from prior assignments.

Classifying Feedback



We used NLP and logistic regression to build a classifier for each feedback criterion. We trained this classifier on manually labeled feedback from three courses.

Dynamic Suggestions Lead to More Benefits than Static

CritiqueKit has been used by 95 reviewers and 336 feedback recipients, in 2 real-world deployments and 2 experiments. We found that linking suggestions dynamically with the real-time checkboxes is more effective at improving feedback than static suggestions. The table to the right summarizes our main findings from these studies.

Study	Adaptive Reuse	Static Reuse	Interactive Guidance	Participants	n=	Main Finding
DEP1		X		TAs	8	TAs used suggestions as inspiration
DEP2		X	X	Design Students	29	Students reused vague suggestions
EXP1		X	X	Design Students	40	Static suggestions and guidance were not helpful
EXP2	X		X	General Students	47	Adaptive suggestions and guidance were helpful