Accuracy and coverage performance profiles across task types and Snuba polarity