VICTORIA UNIVERSITY OF WELLINGTON
Te Whare Wananga o te Upoko o te Ika a Maui

Evidence that the energy detector is not an appropriate model for narrow-band short-duration Gaussian noise detection in humans

Judi A. Lapsley Miller, Vit Drga, & Linton M. Miller
Psychophysics Laboratory, Victoria University of Wellington, New Zealand

Sponsored by ARO member Lynne Marshall
Naval Submarine Medical Research Laboratory, Groton, CT, USA

Abstract

The energy detector is usually proposed as the ideal observer for the detection of narrow-band short-duration Gaussian noise, masked by wider-band Gaussian noise. For some values of the bandwidth-duration (WT) product, however, the energy detector (filter, square-law rectifier, and true integrator) is not as good as a full-linear detector (filter, linear rectifier, and true integrator). Because human decisions in detection tasks are subject to inconsistency, performance rarely approaches that of an ideal observer. This makes it difficult, if not impossible, to evaluate whether the energy detector is appropriate for modeling human detection. By repeating a noise-in-noise detection experiment multiple times using the same stimuli, then using group operating characteristic (GOC) analysis and function-of-replications-combined estimation (FORCE) analysis, the effects of observer inconsistency are minimized and asymptotic error-free performance estimated, respectively. At least for one combination of noise bandwidth and duration, we show that a human observer can do better than the energy detector and do as well as a full-linear detector, once error due to inconsistency is removed. This finding implies that the energy detector is not necessarily an appropriate model for the task of noise-in-noise detection in humans.

Why do humans perform worse than theory in psychophysical tasks?

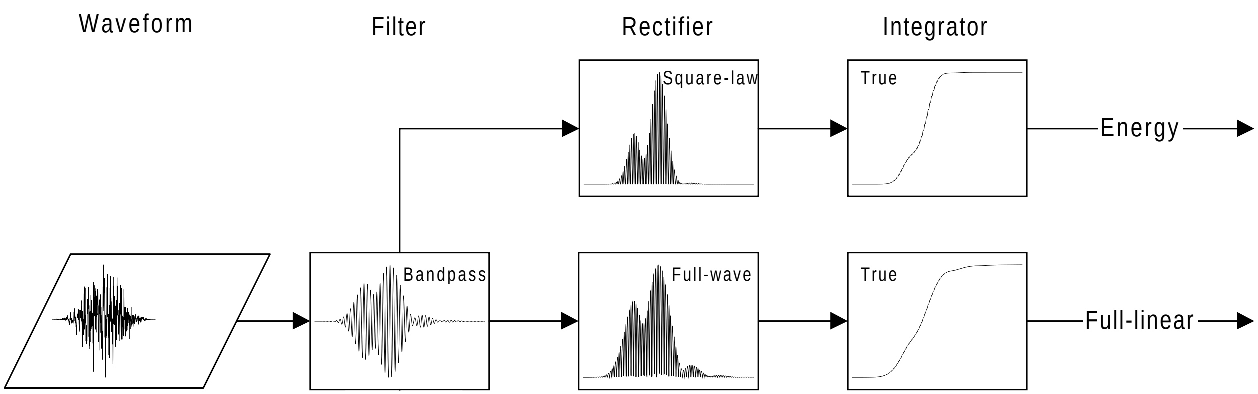
- Mathematical theory may not accurately reflect reality due to simplifications in stimulus representation
- Observer inconsistency when making decisions degrades human performance
- Possible solutions include:
  - Using computer simulations to model theory using more realistic stimuli
  - Reducing or removing observer inconsistency through experimental design and statistical methods

Ideal observers

- Ideal observers are theories that predict the maximum detectability of a class of signals
- Ideal observers are used in psychophysics to
  - put an upper bound on detectabilty
  - find insights into human detectability by studying how it differs from ideal detectability
  - find parsimonious explanations of human detectability
    - it is "unnecessary to invent psychological mechanisms to explain a change that may be traced to the stimulus situation itself" (Green & Swets, 1966).
- Ideal detection $\neq$ perfect detection

Detectability of noise

- The energy detector is normally considered the ideal observer for detecting noise signals masked by wide-band noise (Green, 1960; Green & McGill, 1970)
  - filter
  - square-law rectifier
  - true integrator
  - two parameters
    - signal-to-noise ratio
    - bandwidth-duration product (WT)

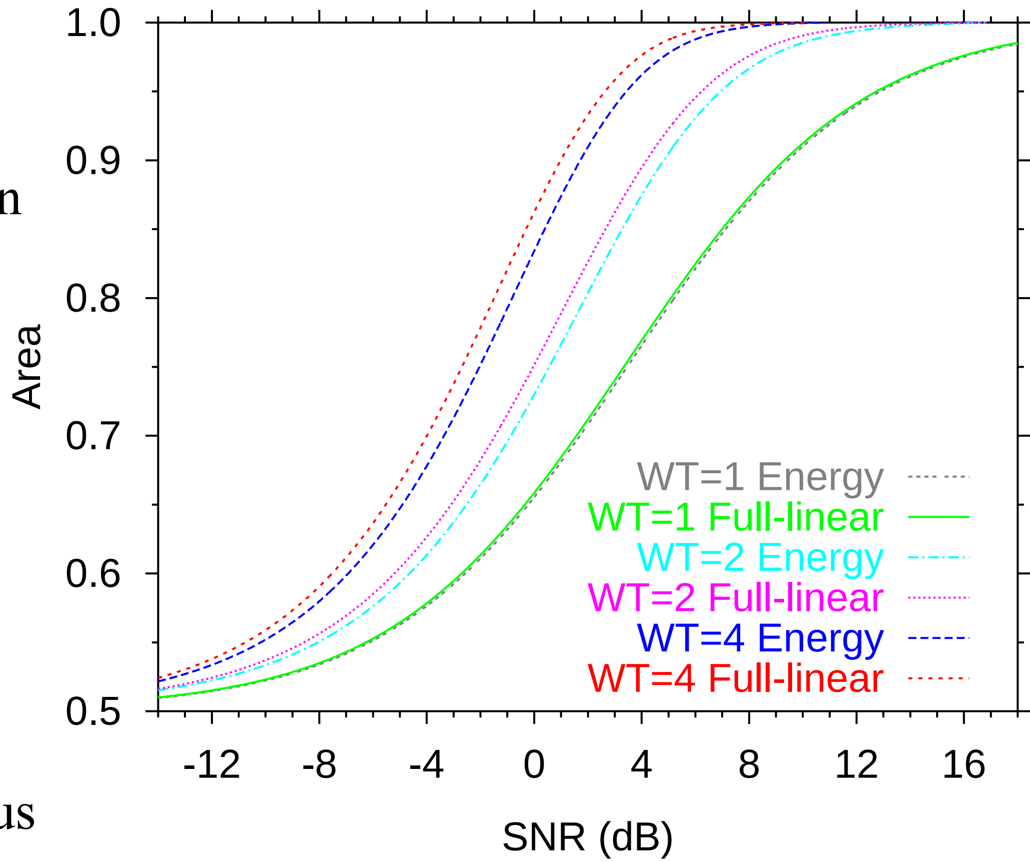


- But a detector using a linear (full or half-wave) rectifier is better than the energy detector for small-WT noise
  - We call this detector a *full-linear detector*
    - There is currently no mathematical theory
    - The engineering literature suggests that the full-linear detector is better then the energy detector for 1<WT<70, the energy detector is better for WT>70, and that they are equally good for WT=1 and 70

Simulated ideal observers

- Stimuli used in mathematical theories are idealized
  - Differences between theoretical and real stimuli are noticeable for Gaussian noise (WT=1, 2, and 4)
  - Psychometric functions for simulated energy detector different to mathematical theory
  - Simulations allow evaluation of theories that may be mathematically intractable
  - Differences between humans and theory may be due to idealized stimulus representation
- Human and simulated detectability can be compared for the same stimuli

Comparison of Psychometric Functions for Full-linear and Energy Detectors



- Simulations show that the full-linear detector is better than the energy detector for WT=2, 4

Observer inconsistency

- Observer inconsistency occurs when an observer's decisions (ratings) differ across repeated presentations of the same stimulus
  - It occurs in all psychophysical tasks
  - It decreases performance and thus detectability
- It is difficult to evaluate psychophysical theory when human performance is degraded by observer inconsistency

Internal & external noise

- Observer inconsistency results from variability (noise) internal and external to the observer
  - Continuous background noise masker
  - Heartbeat, breathing, muscle tension, neural noise
  - Memory, inattention, sequential dependencies, and coordination when using response manipulanda
  - Transient external noise: cars, aircraft, voices
- It is difficult to account for each separate noise source and the effect it has on detectability
- A different approach is to consider the effects of noise, rather than the source

Unique & common noise

- Variability in an observer's ratings can be partitioned into *unique* noise and *common* noise
  - On any one experimental replication, unique and common noise are confounded
  - When an experiment is replicated multiple times, unique noise in an observer's ratings is the component of variability that differs across experimental replications
  - Common noise in ratings is the same across replications and is generally what we are interested in measuring
- Unique noise can be averaged out of ratings, leaving behind common noise

Removing observer inconsistency

- Unique noise can be removed by

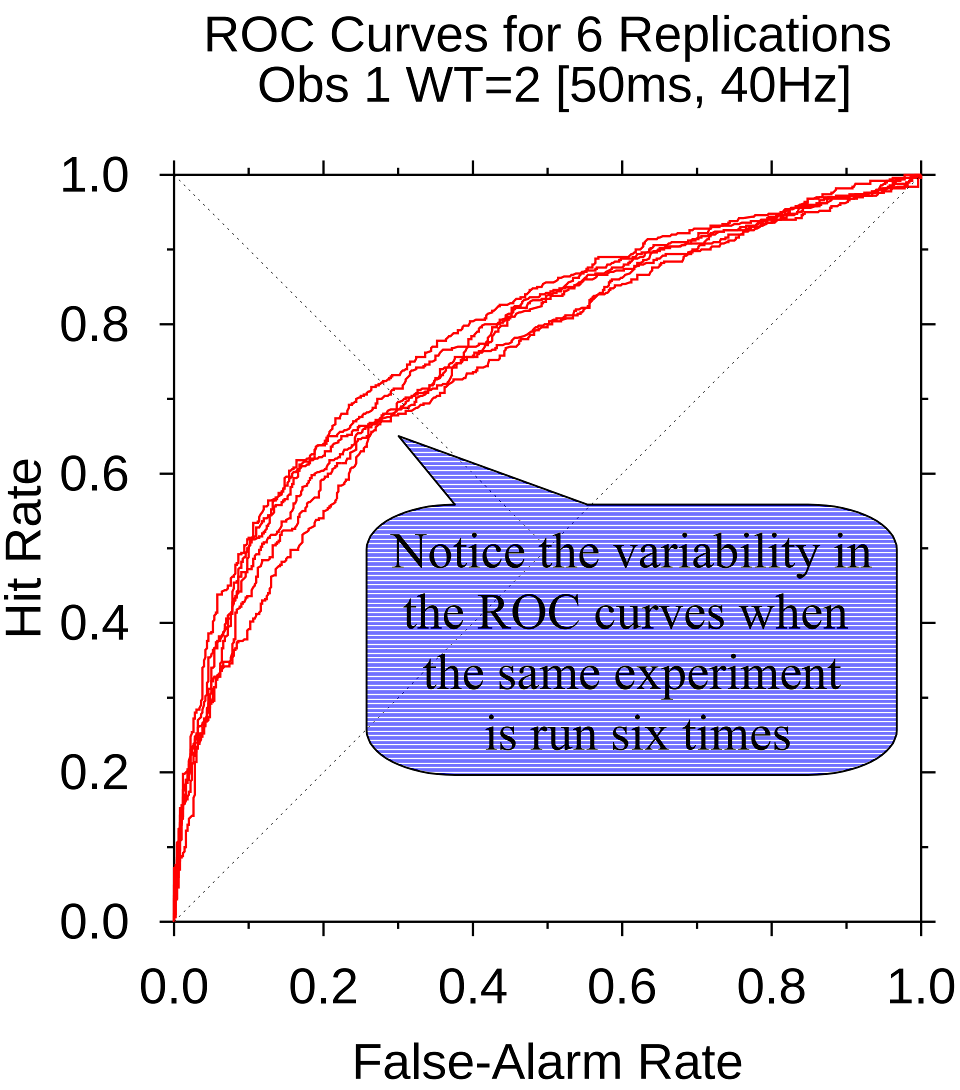
  1) running a multiple-replication experiment

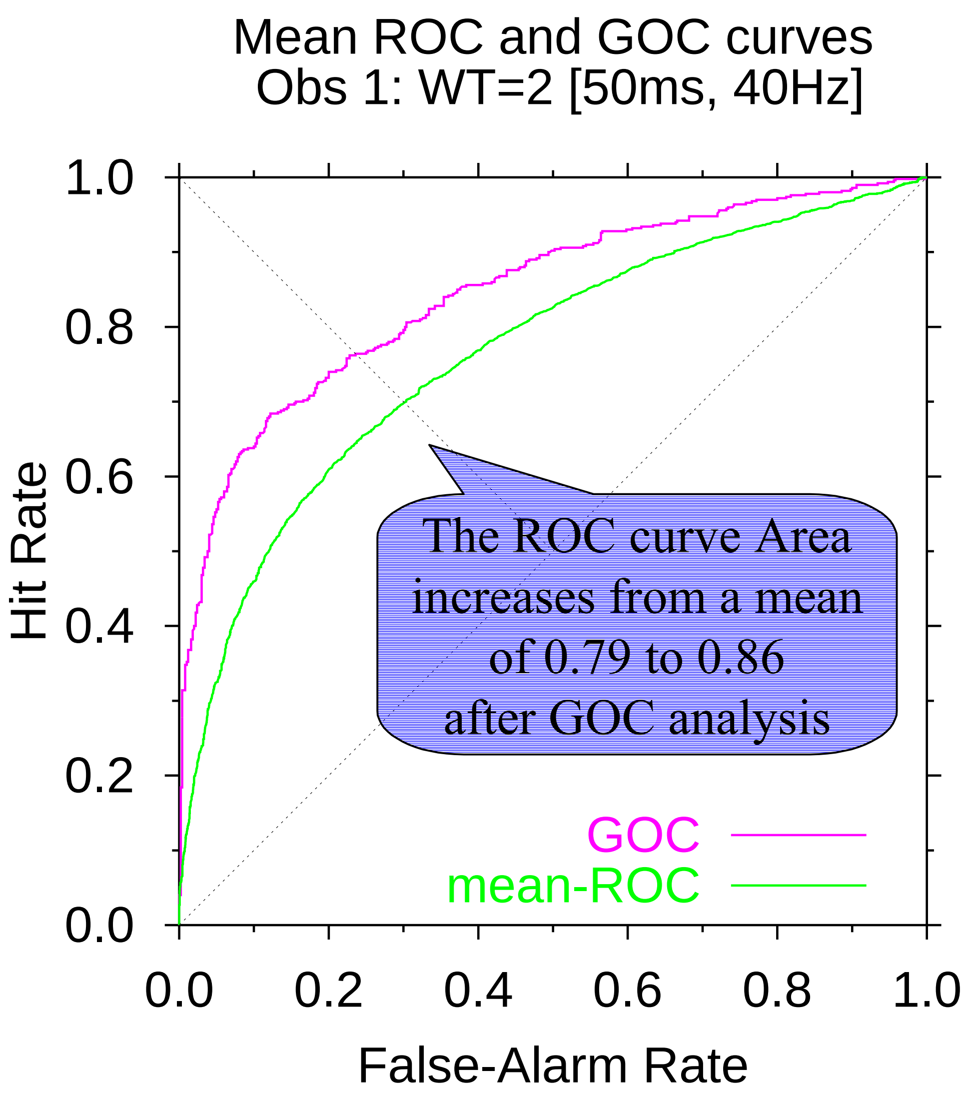
  2) averaging ratings across replications for the same stimulus
- Through this experimental design:
  - Reproducible stimuli are the main sources of common noise
  - Variability from inattention, body noise, etc., are unique noise sources and their effects are averaged out
- Average ratings form the basis for two data analyses
  - Group Operating Characteristic (GOC) analysis, and
  - Function of Replications Combined Estimation (FORCE) analysis

Group Operating Characteristic (GOC) analysis

- Each replication of an experiment gives a single ROC curve–a plot of hit rate against false-alarm rate in a detection task
  - Each single-replication ROC curve is based on one rating per stimulus
- A GOC curve is an ROC curve based on the mean-rating per stimulus, averaged across replications
  - Because GOC removes unique noise, the GOC curve is generally higher than any of the contributing ROC curves
- The next panel (#12) shows
  - ROC curves from six replications of an experiment for one observer in a noise-in-noise detection task
  - The mean-ROC curve and GOC curve based on the same replications

Effect of GOC in ROC space

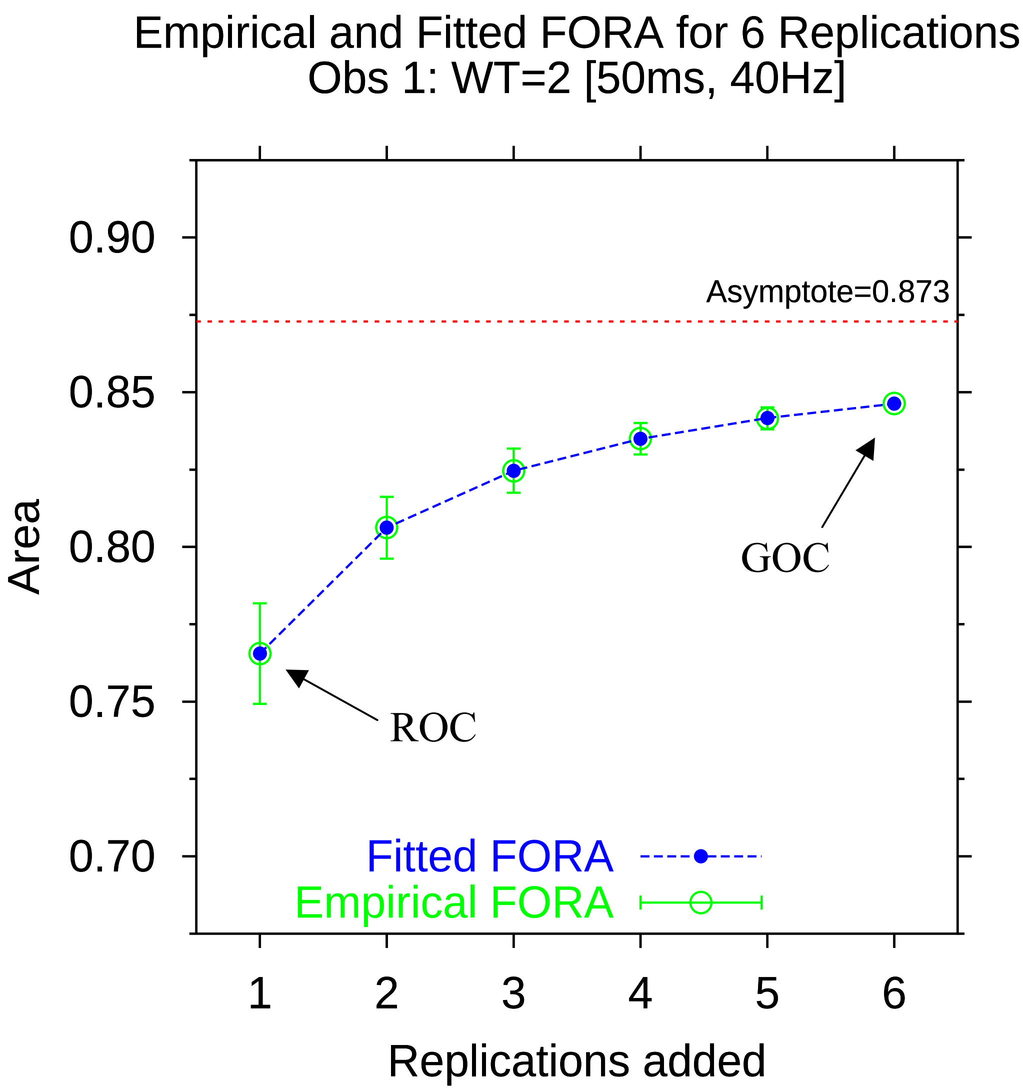




Asymptotic error-free performance

- A function of replications added (FORA) shows how performance improves, on average, as more replications are averaged together in GOC analysis
- FORAs increase smoothly and tend to an asymptote
  - The asymptote indicates the performance level that is possible once all of the error due to observer inconsistency is removed
  - In principle, removing *all* inconsistency requires an infinite number of replications
  - In practice, a particular function can be fitted to a FORA (from a finite number of replications) and extrapolated to infinity to estimate the asymptote
- Fitting this function and extrapolating to the asymptote is called function of replications combined estimation (FORCE) analysis

Example of FORCE analysis

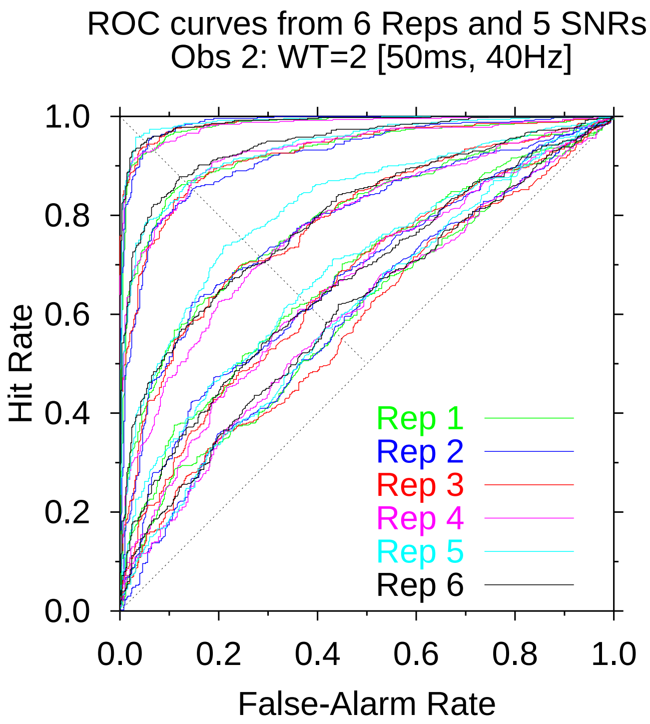


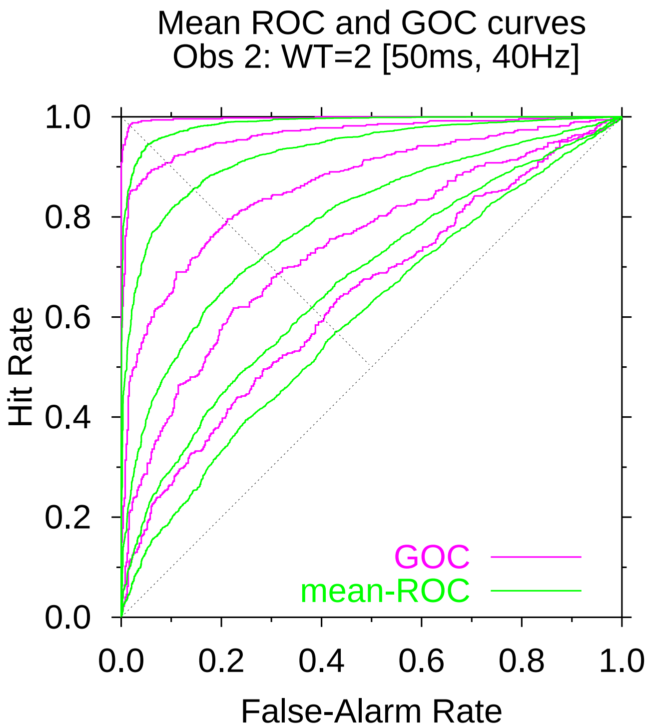
- This is the empirical and fitted FORA for the data set from panel #12
  - The measure of performance used here is Area under the GOC curve
  - FORCE works for any measure of sensitivity
- The first point on the empirical FORA shows the average single-replication ROC performance
- The last point on the empirical FORA shows GOC performance after six replications
- The asymptote represents unique-noise free performance
  - In this case, detectability would improve further if more replications were run
- Human asymptotic detectability can be compared to an ideal observer without an attenuation factor

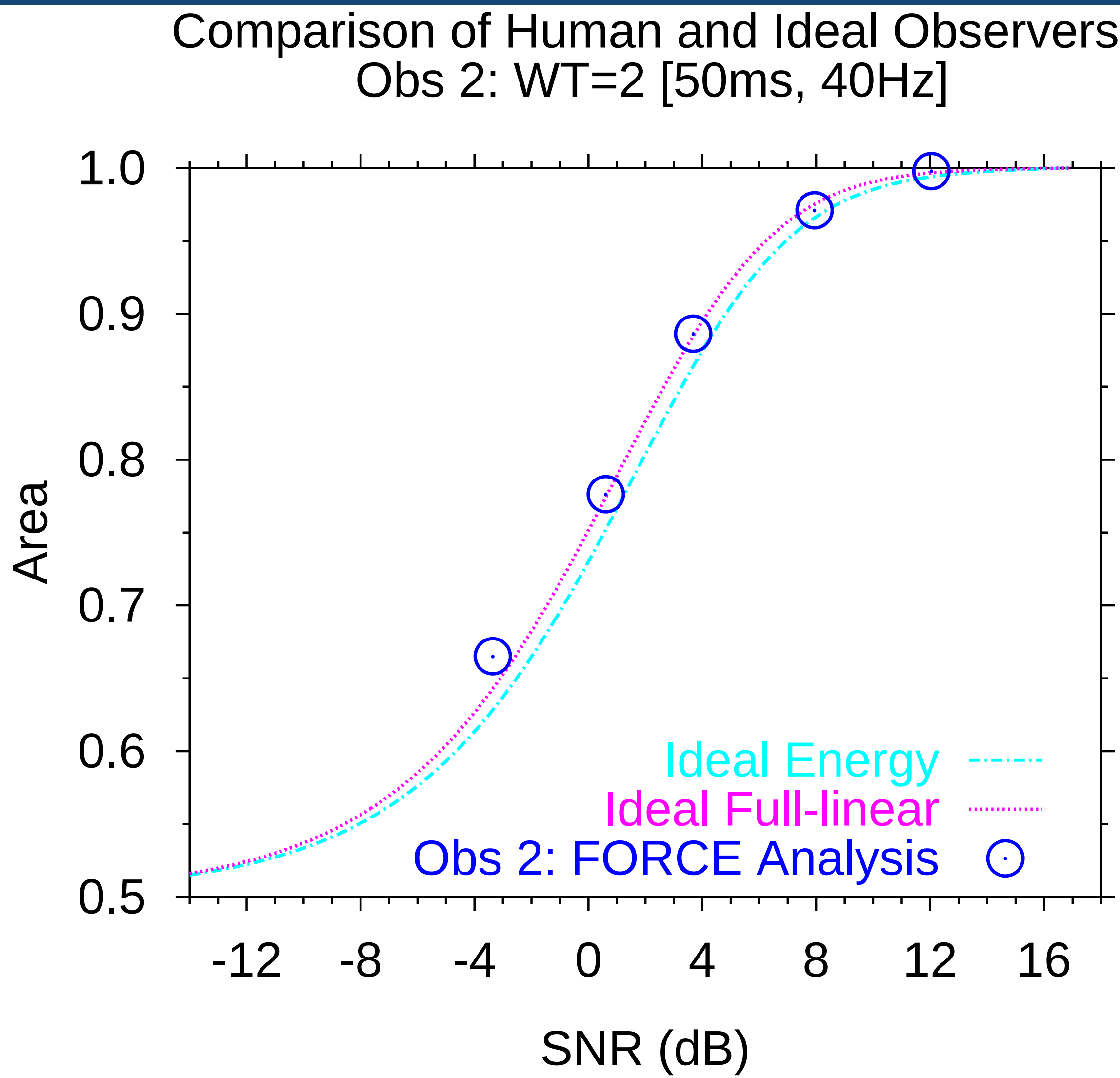
Human experiment repeated 6 times

- Task was to detect small-WT Gaussian noise signals
  - Single-interval forced-choice task
  - WT=1, 2, 4
    - W: 2.5 to 160 Hz
    - T: 6.25 to 400 ms
  - 18 W&T combinations
  - 5 signal-to-noise ratios
  - 500 stimuli per event
  - Wide-band noise masker
    - 4000 Hz, low pass
  - Two observers
    - Repeated experiment 6 times
    - 324000 trials per observer
- ROC analysis
  - lots of variability over the six replications
- GOC analysis
  - always improved performance
    - 2 to 3 dB from ROC
- FORCE analysis
  - generally improved performance further
    - 0.5 to 2 dB from GOC
- For some W&T combinations
  - Performance equivalent to ideal, simulated, full-linear detector

Human does better than energy detector





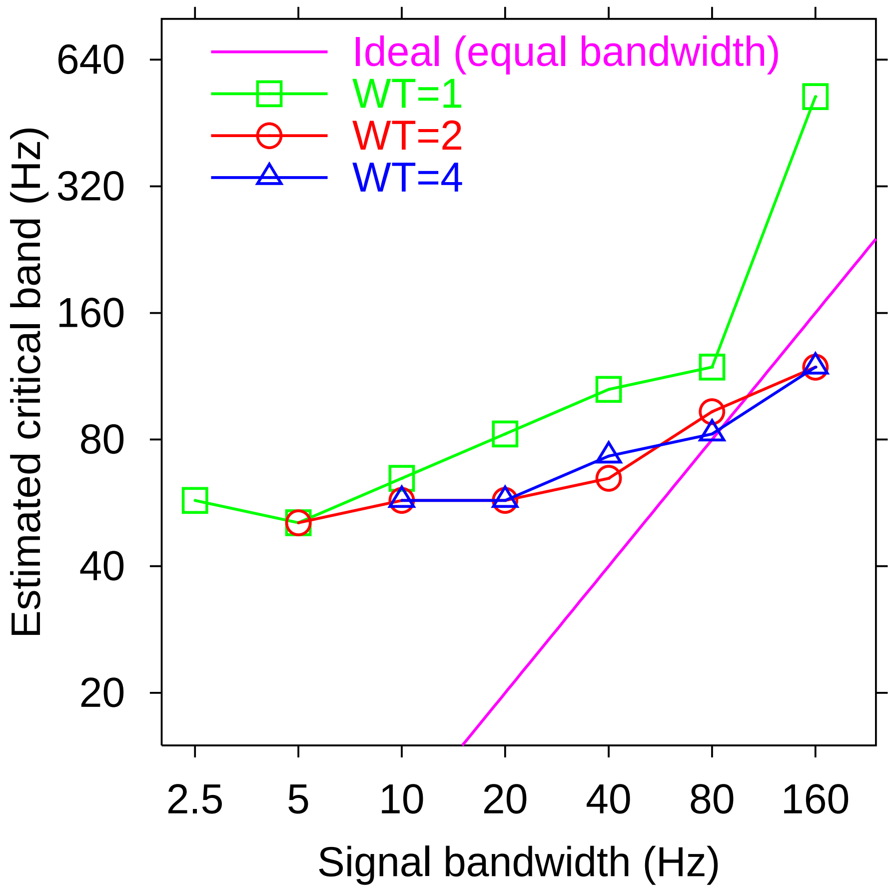


Interpretation of results

- In all cases, GOC analysis improved performance as replications were added
- Humans were better than a simulated energy detector for some WT combinations
- Human performance was as good as an ideal simulated full-linear detector for some WT combinations
- For most WT combinations humans did worse than ideal even after FORCE analysis
  - This difference is NOT due to observer inconsistency
  - Perhaps the human auditory filter (critical) bandwidth was wider than the signal bandwidth?
  - Could human detectability still be like a full-linear detector, but with a sub-optimal filter bandwidth?

Human & simulation correlations

- Correlations were calculated between human and simulation ratings for the same stimuli, over many simulation filter bandwidths
  - In general, correlations were larger for the full-linear detector, compared to the energy detector, regardless of bandwidth
- Graph shows how the best-correlated simulation filter bandwidth varied with signal bandwidth
  - Auditory filter (critical) bandwidths are for Obs. 2, estimated from the sub-optimal full-linear detector simulations
  - Best-correlated bandwidth was not constant, nor was it ideal



Summary

- A human was shown to perform like a full-linear detector, and better than an energy detector in a noise-in-noise detection task
- Computer simulations showed that the energy detector is not as good as a full-linear detector for detecting Gaussian noise with WT=2 and 4
- GOC and FORCE analyses remove the effects of observer inconsistency
  - When observer inconsistency is reduced, human detectability can be better compared to theoretical detectability
  - Asymptotic error-free performance may be estimated from as few as six replications (although more replications result in more reliable estimates)

For further information...

- Available at www.psychophysics.org
  - Drga, V. (1999) *The theory of group operating characteristic analysis in discrimination tasks*. PhD Thesis, Victoria University of Wellington, New Zealand.
    - Theory of GOC analysis and development of FORCE analysis
  - Lapsley Miller, J. A. (1999) *The role of the bandwidth-duration product in the detectability of diotic signals*. PhD Thesis, Victoria University of Wellington, New Zealand.
    - Human experiments and simulated ideal observers
  - Lapsley Miller, J. A. (2000) New techniques to reduce the effects of observer inconsistency in psychoacoustic experiments. *Journal of the Acoustical Society of America*, 107 (5) 2914(A). (Full poster available)
    - Tutorial-style overview of GOC and FORCE with examples from various experiments. Gives simple example of GOC calculations.
  - Linton Miller's C code for FORCE analysis
- Available in your local library

  - Taylor, A., Boven, R., and Whitmore, J. (1991). Reduction of Unique Noise in the Psychophysics of Hearing by Group Operating Characteristic Analysis. *Psychological Bulletin, 109*(1), 133-146.
  - Green, D. (1960). Auditory Detection of a Noise Signal. *J. Acoust. Soc. Am., 32*, 121-131.
  - Green, D. & Swets, J. (1966). *Signal Detection Theory and Psychophysics*. Wiley: NY.
  - Green, D. & McGill, W. (1970). On the equivalence of detection probabilities and well-known statistical quantities. *Psych. Rev., 77*, 294-301.

- Contact Details
  - Victoria University's Psychophysics Lab closed in February 2000 when John Whitmore retired. The lab continues to exist in a more abstract sense on the internet at www.psychophysics.org. The authors may also be contacted at their current affiliations:

- Judi Lapsley Miller
  - NSMRL, Subase NLON, Groton CT 06349, USA. ph 860-694-2557.
  - email jmiller@nsmrl.navy.mil or judi@psychophysics.org
- Vit Drga
  - Dept. of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom
  - email vitdrga@essex.ac.uk or vit@psychophysics.org
- Linton Miller: c/- Judi Lapsley Miller
  - email linton@psychophysics.org