Inhibitory tagging in visual search: A failure to replicate

JEREMY M. WOLFE and CHARLES W. POKORNY
Massachusetts Institute of Technology, Cambridge, Massachusetts

Klein (1988) reported that increased reaction times for the detection of small light probes could be used as an indicator of inhibitory tagging of rejected distractors in serial visual search tasks. Such a paradigm would be very useful in the study of the mechanics of visual search. Unfortunately, we cannot replicate the result. In this study, we found that probe reaction times were elevated at *all* distractor locations, relative to empty locations, following both parallel and serial search tasks. This appears to be a forward masking effect.

Many searches for a target item among distractor items appear to require the deployment of attention from item to item in a serial fashion (e.g., Treisman & Gelade, 1980). Once a distractor has been visited by attention and rejected as a candidate target, it would seem reasonable to prevent attention from paying a return visit. Posner and his colleagues (Posner & Cohen, 1984; Posner & Presti, 1987) have found that it is more difficult to direct attention to a recently attended location than to an unattended location. Reaction times (RTs) for small light probes are longer at this inhibited location. They have called this phenomenon "inhibition of return." Klein (1988) applied Posner's concepts and methods directly to the visual search paradigm. He reasoned that inhibitory tagging of distractors might prevent attention from returning to rejected distractors and that this inhibition might be measurable with Posner's probe-detection paradigm. Briefly, RTs should be longer at rejected distractor locations than at empty locations in the field. The effect should occur for serial search tasks but not for parallel search tasks. Finally, it should be about twice as great for target-absent search trials as for target-present trials, since all distractors should be inhibited on target-absent trials and an average of only half of the distractors should be inhibited in target-present trials. Klein's results supported these conjectures. In a serial search task, probe-detection RTs (which we will refer to as probe RTs) were 50-60 msec longer for probes placed in rejected distractor locations than for probes placed in empty locations in the field. The effect was not seen for a parallel search task and was approximately half as large on target-present trials.

This paradigm could be exceedingly useful in the study of the details of visual search behavior. Averaging over trials, it could be used to "count" the number of inhibited distractors in a given condition. Klein demonstrates this potential in his experiments. In a serial, self-terminating search, attention should visit and reject all distractors when there is no target present, but only half of the distractors when a target is present. Given 50% inhibited and 50% uninhibited locations on target trials, the average magnitude of the RT increase for target trials should be half that for target-absent trials. Klein reports just such results.

Recently, Wolfe, Cave, and Franzel (1989) proposed that some searches do not require full serial, self-terminating search because the parallel, preattentive stage of processing can guide attention toward likely targets. Thus, in a search for a red T among green Ls and red Ls, even though the search for a T among Ls requires attention, the effective number of distractors can be cut in half by guiding attention toward the red items and away from green (see also Egeth, Virzi, & Garbart, 1984). Evidence for this selective, or guided, search could be provided using Klein's paradigm if it could be shown that on target-absent trials all red locations were inhibited while all green locations were not. Similar experiments could be done to test similar hypotheses about performance on a variety of other search tasks. We thus set out to replicate and extend Klein's findings. It is with considerable regret that we report that we have been unable to do so.

EXPERIMENT 1

Method

Apparatus and Procedure. The stimuli for Experiment 1 were displayed on a standard television monitor that was part of a modified "Sub Roc 3-D" video game. The display was controlled by an IBM PC-XT with IBM YODA graphics. The stimuli were presented in an 11.3° × 11.3° field bounded by a green line. Individual stimuli were constrained to fit within a 0.85° × 0.85° field. The stimuli were positioned on an invisible, slightly irregular 6 × 6 array. A fixation spot was provided at the center of the display, and the subjects were instructed to fixate. In Experiment 1, stimuli closely followed those of Klein (1988). Stimuli were circles with or without

a small vertical line that intersected the circle at "12 o'clock" (i.e., O and Ô). Phosphor decay on this device is quite slow.

The subjects were tested twice in the standard search paradigm without probes and twice in the Klein probe procedure as described below. In each case, one block of trials used the Ô as the target and the Os as distractors. In the other block, target and distractors were reversed. Search for an O among Ôs is *serial*, whereas search for a Ô among Os is *parallel* (or, at least, more efficient than the search for O among Ôs; Treisman & Souther, 1985). The O-among-Ôs task should require the deployment of attention from item to item and, thus, should yield evidence for inhibition of return in the form of elevated RTs for probes presented in distractor locations. The Ô-among-Os task should not show evidence of this inhibition.

In the search task, the subject initiated each trial by pressing a button with the left hand. This caused a search display to appear. On 50% of the trials, a target was present among distractors. On the remaining trials, only distractors were present. The subject responded by pressing one of two buttons with the right hand, depending on whether or not the target was present. The display remained on the screen until the subject responded. RT was measured from stimulus onset to the response of the subject Feedback was given after every trial. Placement of target and distractors was random within the field described above. Set sizes of 5, 10, and 20 items were used and varied randomly from trial to trial. The subjects were tested for 30 practice and 300 "real" trials.

Two variants of the probe version of the experiment were run. For both, set size was fixed at 10 items and the initial search portion of each trial was identical to that described above. In Variant 1, a small (4 pixel), 6.5 cd/m² probe was presented 60 msec after the offset of the search stimulus on 75% of the trials (ISI taken from Klein,1988). In Variant 2, the probe was presented on 50% of the trials, exactly following Klein's methods. For both variants, on half of these probe-present trials, the probe was presented at a distractor location. On the remaining trials, the probe was presented in an empty location Probes were placed in the black center of the O or Ô or in a black empty location. The probe remained visible until the subject responded or the trial terminated. Fairly

dim search stimuli were used in an effort to minimize forward masking effects (lum: 1.1 cd/m²). Probe location and trial type were random across trials. Responses were slightly different in the two variants. In Variant 1, the subjects responded to the presence of the probe by pressing either of the right-hand buttons, the same buttons used to respond in the search portion of the task. In Variant 2, the subjects responded with the left hand. This use of different hands for the two parts of the task follows Klein's methods. The subjects withheld response in the absence of probe. A trial without a probe terminated after 1,500 msec. Feedback was provided after each trial. The subjects were tested for 50 practice and 250 real trials in the probe condition. Variants 1 and 2 were similar, with Variant 2 being the more faithful replication of Klein's methods.

In the search component, trials having RTs greater than 5,000 msec and less than 250 msec were discarded and were not replaced. On the probe component, RTs of less that 150 msec were discarded. Fewer than 1% of trials were discarded for these reasons. RTs of greater than 1,500 msec on a probe-present trial were considered to be missed probes.

Subjects. There were 10 different subjects in each variant. The subjects were either employees of our lab or were obtained from the MIT subject pool. The subjects from the MIT subject pool gave informed consent prior to the experiment and were paid $6/h. The subjects' ages ranged from 18 to 24 years. The subjects wore their best optical correction.

Results

RTs for the search tasks were averaged across subjects. These means, plotted as a function of set size, are shown in Figure 1 for target-present and target-absent trials separately. Slopes of best-fit linear regressions are also shown on the graph. Error rates in the search task averaged less than 3%. An ANOVA yielded significant effects of set size and trial type (target present vs. target absent) for both the Ô-among-Os and the O-among-Ôs conditions (all
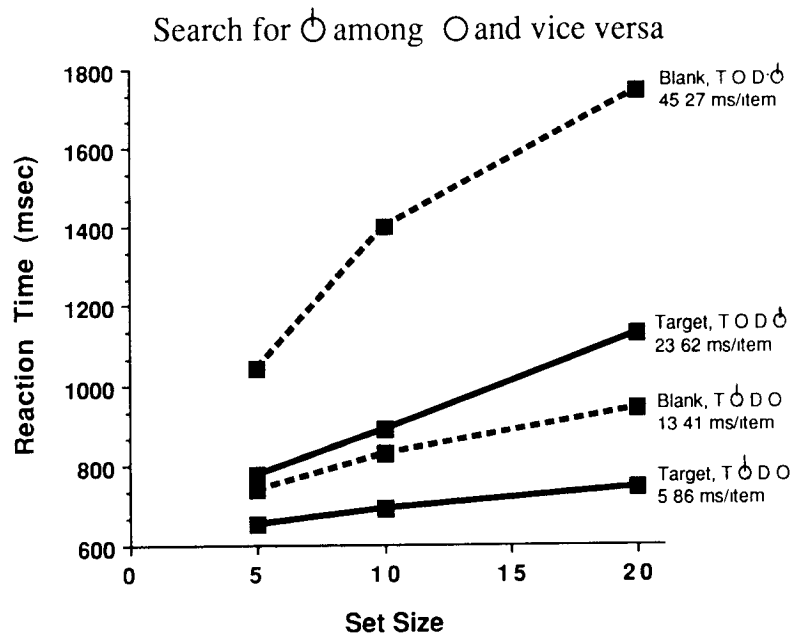

Figure 1. Reaction time as a function of set size for target-present and target-absent (blank) trials in two different search tasks. Note that it is much easier to find a Ô among Os than an O among Ôs.

$p < 0.005$). The slopes for the O-among-Ȯs condition were significantly steeper than those for the Ȯ-among-Os condition for target-present trials [$t(9) = 4.95, p < .0005$] and target-absent trials [$t(9) = 6.17, p < .0005$).

There are four probe conditions of interest: The probe could be presented on a target-present or a target-absent trial and could be presented in a distractor location or in an empty location. The RT values for these four conditions, averaged across subjects, and the difference between distractor and empty location probe RTs are shown in Table 1 for both variants.

In Variant 1, the differences between distractor and empty locations were significant for all conditions [all ts(9) > 5.0, all ps $< .001$]. None of the comparisons between target-present and target-absent search trial conditions revealed significant differences in probe RTs [all ts(9) < 1.3, all ps $> .1$]. None of the comparisons between parallel and serial search tasks reveal significant differences in probe RTs [all ts(9) < 1.4, all ps $> .1$]. In Variant 2, the differences between distractor and empty locations were significant for all conditions [all ts(9) > 3.3, all ps $< .01$]. Probe RTs were significantly longer for probes presented in empty locations on target-absent trials than for those presented in empty locations on target-present trials [$t(9) = 4.9, p < .001$]. None of the other comparisons between target-present and target-absent search trial conditions revealed significant differences in probe RTs [all ts(9) < 1.6, all ps $> .1$]. None of the comparisons between parallel and serial search tasks revealed significant differences in probe RTs [all ts(9) < 0.8, all ps $> .4$]. There was a larger effect of probe location for target-present trials than for target-absent trials in the serial task [$t(9) = 3.4, p < .01$]. This effect was not significant for the parallel task [$t(9) = 1.1, p > .2$].

Error rates. In Variant 1, error rates for the search component of the task were 2.6% for the O-among-Ȯs task and 2.7% for the Ȯ-among-Os task. For the probe component, the rates were 3.5% for the O-among-Ȯs task and 3.1% for the Ȯ-among-Os task. In Variant 2, error rates for the search component of the task were 4.3% for the O-among-Ȯs task and 2.0% for the Ȯ-among-Os task. For the probe component, the rates were 1.6% for the O-among-Ȯs task and 2.3% for the Ȯ-among-Os task.

Discussion

If probe RTs can be used to measure inhibition of return, the pattern of results should be as follows: RTs for probes presented in distractor locations should be longer than those presented in empty locations. This effect should only appear in the serial search task since the parallel search task presumably does not involve inhibition of individual distractor locations. The effect should be about twice as large for target-absent search trials as for target-present search trials because, on target-absent search trials, all distractor locations should be inhibited whereas, on target-present search trials, averaged across trials, only half of the distractor locations should be inhibited. In Experiment 1, probe RTs were longer when probes were presented in distractor locations, as reported by Klein (1988). However, experiment type and trial type had no influence on this effect. It is the prior presence of a distractor that seems to make the probe more difficult to detect. This may be considered to be a form of forward masking (see Breitmeyer, 1984, for a comprehensive review). The one exception to this conclusion is the significant effect of search trial type on size of the "inhibition" effect in serial block of Variant 2. However, the difference goes in the wrong direction. *More* inhibition was seen on target-present trials, when the average inhibition on those trials should have been half of that seen on target-absent trials.

The variations of percentage of probe trials and hand of response that differentiate the two variants of Experiment 1 do not appear to have had any important impact on the pattern of results. Neither variant provides evidence for inhibitory tagging. One further variable differentiates our methods from Klein's. Our stimuli were presented in an $11.3° \times 11.3°$ field. In Klein's experiment, the display size was $5° \times 6°$. It is hard to imagine that this variation is critical. If there was any effect, one might expect inhibitory tagging to be easier in our less dense display.

Some of our subjects produced quite "serial" results on the Ȯ-among Os (parallel) task. Perhaps they show inhibitory tagging for both Ȯ-among-Os and O-among-Ȯs tasks, hiding the difference between serial and parallel search. If this is the case, there should be a correlation between slopes on the Ȯ-among-Os search task and the

Table 1
Probe Reaction Times for Ȯ versus O Experiments

Target	O-Among-Ȯs Tasks			Ȯ-Among-Os Tasks		
	Probe Location			Probe Location		
	Distractor	Empty	Difference	Distractor	Empty	Difference
Variant 1						
Present	440	381	59	421	355	66
Absent	448	393	55	423	360	63
Variant 2						
Present	504	368	135	498	370	128
Absent	512	434	78	502	405	96

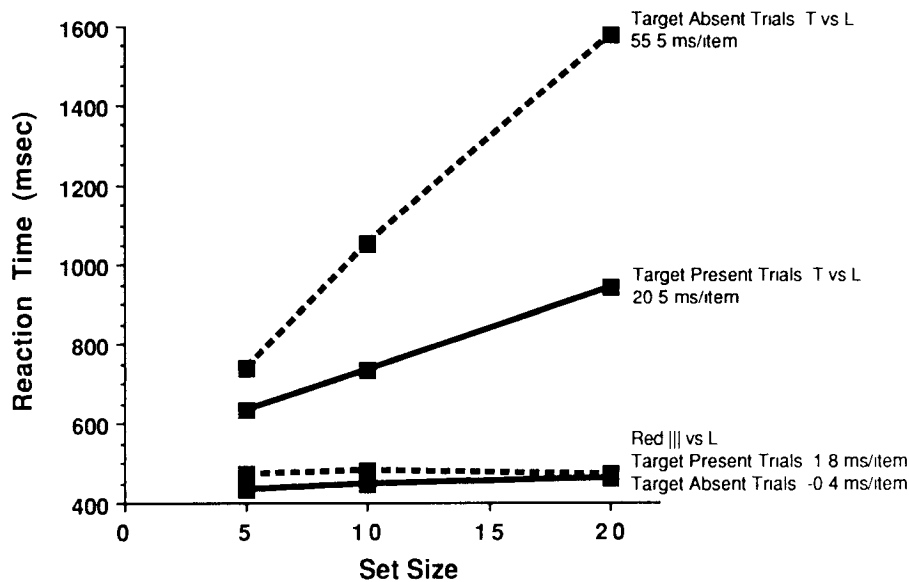Note—All reaction time values are given in milliseconds.

Figure 2. Reaction time as a function of set size for target-present and target-absent (blank) trials in two different search tasks. Note that the T versus L task produces classic *serial* search results, whereas the red grating versus L task produces classic *parallel* results.

difference between distractor and empty location probe RTs. The subjects with high search slopes should have larger RT differences, greater evidence for inhibitory tagging. No such correlation is seen (in Variant 1, $r = 0.148$, $df = 9$).

It might be argued that the serial task is not typical because it is the search for the absence of a basic feature. Similarly, as noted above, it could be argued that the parallel task was not truly parallel because the data show a dependence of search RT on set size. Of course, both of these objections hold for the original Klein study, too. Nevertheless, it seemed worth repeating the experiment with more conventionally serial and parallel tasks.

EXPERIMENT 2

Method

Procedure. As in Experiment 1, two search tasks were used. For both conditions, the distractors were gray Ls in four possible orientations (0°, 90°, 180°, and 270°). The Ls consisted of two sides of a square 0.85° on a side. The probe, when presented at a distractor location, was always presented at the center of that square. The target for the parallel condition of Experiment 2 was a red grating—different in color, luminance, and form from the distractors. The target for the serial condition was a gray T. It, like the

Ls, could appear in four 90° rotations. Search for a T among Ls has been shown to be serial using these stimuli in this lab (Wolfe et al., 1989). All other aspects of the experimental apparatus and methods were identical to those of Experiment 1. As before, two versions were run for each task. In one case, each trial contained only the search task. In the other, each trial was composed of both search and probe-detection tasks.

Subjects. There were 10 subjects for Experiment 2 (6 male and 4 female). The subjects were either employees at the lab or were obtained from the MIT subject pool. The subjects from the MIT subject pool gave informed consent and were paid $6/h. The subjects' ages were between 18 and 23 years. The subjects wore their best optical correction.

Results

Average RTs as a function of set size for target-present and target-absent trials are shown in Figure 2. They show the expected results. Search for a T among Ls was serial, with slopes of standard steepness and a rough 2:1 ratio between target-absent and target-present slopes. The search RT for a red vertical grating among Ls was fast and independent of set size, as would be expected in a parallel search.

Table 2 shows the mean probe RTs for the probe conditions of Experiment 2. RTs for probes presented in distractor locations were significantly longer than those for

Table 2
Probe Reaction Times for Experiment 2

| Target | T-Among-Ls Tasks | | | Red-llll-Among-Ls Tasks | | |
| | Probe Location | | | Probe Location | | |
	Distractor	Empty	Difference	Distractor	Empty	Difference
Present	342	303	39	332	296	36
Absent	344	321	23	319	293	26

Note—All reaction time values are given in milliseconds.

probes presented at empty locations [$t(9) > 2.2, p < .05$, for all four differences]. There were no significant differences in probe RTs between target-present and target-absent trials for either the serial task [$t(9) = 1.3, p > .1$] or the parallel task [$t(9) = 0.5, p > .1$]. There were no significant differences in probe RTs between the serial and parallel tasks [target-present trials, $t(9) = 0.3, p > .1$; target-absent trials, $t(9) = 0.2, p > .1$]. The distractor-empty location differences appear to have been larger for target-absent trials than for target-present trials. This trend was not significant for either task [$t(9) < 1.2, p > .25$].

Discussion

The results of Experiment 2 mirror those of Experiment 1. Probe RTs were reliably longer when probes were presented at distractor locations than when probes were presented at empty locations. Serial and parallel tasks produced a difference of the same magnitude. In addition, target-present and target-absent trials produced similar results. This supports the hypothesis that the effects on probe RT are a form of forward masking. The difference between probe RTs at distractor and empty locations was smaller in Experiment 2 than in Experiment 1 [collapsing across all conditions, $t(78) = 4.5, p < .001$]. If we assume that the effect is a forward masking effect, that would suggest that there was less masking in Experiment 2 than in Experiment 1, perhaps because the distractors in Experiment 1 completely surrounded the probe location, whereas the distractors in Experiment 2 flanked the probe location on only two sides. We tested this possibility in Experiment 3.

EXPERIMENT 3

Method

Procedure and Subjects. In Experiment 3, the search task was a very easy parallel task. The target was a red vertical grating. The distractors were a gray O (from Experiment 1) and a gray L (upright only, from Experiment 2). Probes were presented on 75% of the trials. On probe trials, the probe was presented with equal probability at an O location, an L location, or an empty location. This allowed for direct comparison of the effects of each type of distractor on probe RTs. Eleven subjects were tested. In all other respects, the methods were identical to the probe versions of Experiments 1 and 2.

Results

Table 3 gives the average probe RTs for probes presented at O, L, or empty locations. An ANOVA yielded a significant effect of probe location [$F(2,60) =$

Table 3
Probe Reaction Times for Experiment 3

Target	Probe Location		
	O Distractor	L Distractor	Empty
Present	378	336	314
Absent	387	350	303

Note—All reaction time values are given in milliseconds.

$3.5, p < .05$] and no significant effect of search trial type [$F(1,60) = 0.029, p = .87$]. The interaction was not significant [$F(2,60) = 0.115, p = .89$]. Paired t-test comparisons revealed a significant difference between probes presented at the two types of distractor location [target-present trials, $t(10) = 2.6, p < .02$; target-absent trials, $t(10) = 2.8, p < .01$]. None of the differences between target-present and target-absent search trial conditions were significant [all $ts(10) < 1.3$, all $ps > .1$].

Discussion

There was a significant elevation of probe RTs in this parallel search task. It was different for the two types of distractor. The magnitude of the probe RT effect for probes presented at L locations was comparable to that seen in Experiment 2 where Ls were the distractors. Similarly, the magnitude of the probe RT effect at O locations was comparable to that seen in Experiment 1 where Os were the distractors. There is no obvious reason to suppose that more inhibition should be present at the O locations than at the L locations. These data are consistent with the hypothesis that probe RT effects are the result of forward masking of the probe by the distractor. An O that completely surrounds the location of the subsequent probe flash appears to produce more masking than does an L that flanks the probe location on only two sides. One could study the parameters of this phenomenon as a way to confirm the hypothesis, although that does not seem very interesting.

We are quite unhappy with our results. Klein's (1988) paper seemed to offer a fairly easy and very useful paradigm for the investigation of the deployment of attention in visual search. It is not clear why we did not obtain Klein's reported pattern of results. Our probe RT effects are of about the same magnitude as his, making it unlikely that noise in our data was hiding the inhibitory tagging effect. We have attempted a fairly close replication of his study, although there were some differences in methods (as discussed in Experiment 1). Perhaps such differences in the experimental parameters do account for the differences in our results. If so, the nature of the critical variable becomes important. If the Klein paradigm works only for a very limited set of search tasks, then it is of little general value. However, if there is some inconsequential change in the probe task that can make the paradigm work reliably, then it may yet prove useful. Unfortunately, our informal efforts to vary probe luminance, size, and ISI did not reveal any condition that produced Klein's pattern of results.

In summary, in our hands, the probe-detection paradigm does not yield evidence for inhibitory tagging of distractor locations in visual search. We find that RTs for detection of probes are longer at distractor locations than at empty locations, but this effect is seen for both parallel and serial search tasks and is of the same magnitude for probes presented after target-present or target-absent search trials. While we do not know the source of the difference

between our results and those of Klein (1988), we suspect that the longer probe RTs at distractor locations in our experiments are the result of forward masking.

REFERENCES

BREITMEYER, B. G (1984) Visual masking. An integrative approach Oxford. Oxford University Press.

EGETH, H E , VIRZI, R A , & GARBART, H. (1984) Searching for conjunctively defined targets Journal of Experimental Psychology Human Perception & Performance, 10, 32-39

KLEIN, R. (1988). Inhibitory tagging system facilitates visual search. Nature, 334, 430-431

POSNER, M I., & COHEN, Y (1984) Components of attention In H Bouma & D. Bowhuis (Eds), Attention and performance Hillside, NJ: Erlbaum.

POSNER, M I , & PRESTI, D E (1987) Selective attention and cognitive control Trends in Neuroscience, 10, 13-17.

TREISMAN, A M , & GELADE, G. (1980). A feature integration theory of perception Cognitive Psychology, 12, 97-136

TREISMAN, A M., & SOUTHER, J (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. Journal of Experimental Psychology· General, 114, 285-310.

WOLFE, J. M., CAVE, K. R , & FRANZEL, S. L. (1989) Guided search. An alternative to the feature integration model for visual search Journal of Experimental Psychology: Human Perception & Performance, 15, 419-433

NOTE ADDED IN PROOF

Since this paper was accepted, we have become aware of three other unsuccessful efforts to duplicate or extend Klein's results (A. Treisman, J Zacks, R Ivry, and A. Cohen, personal communications)

Notices and Announcements

20th Annual Meeting of the Society for Computers in Psychology
New Orleans, Louisiana
November 15, 1990

The 20th Annual Meeting of the Society for Computers in Psychology will be held at the Hyatt Regency Hotel in New Orleans on November 15, 1990. The meeting will include presentations, discussions, and times to preview software and hardware. All areas of psychology are featured, including research, education, clinical practice, and industrial applications.

For further information, contact Sarah Ransdell, Department of Psychology, University of Maine, Orono, ME 04469 (BITNET address RANSDELL@MAINE) or C. Michael Levy, Department of Psychology, University of Florida, Gainesville, FL 32611 (BITNET address MLEVY@UFFSC).

31st Annual Meeting of the Psychonomic Society
New Orleans, Louisiana
November 16-18, 1990

The 31st Annual Meeting of the Psychonomic Society will be held in New Orleans, November 16-18, 1990. The meetings will begin Friday morning and continue until Sunday at noon. The headquarters hotel will be the Hyatt Regency.

A copy of the program will be published in the November issue of the Bulletin of the Psychonomic Society.

For further information, please contact the secretary-treasurer of the Society: Cynthia H. Null, Psychology Department, College of William & Mary, Williamsburg, Virginia 23185 (Telephone: 804-221-3882; BITNET WPSLCHN@WMMVS).