Studies in Avian Biology No. 26:177–179, 2003.

A FIELD EVALUATION OF THE SOUTHWESTERN WILLOW FLYCATCHER SURVEY PROTOCOL

ROLAND S. SHOOK, SCOTT H. STOLESON, AND PAUL BOUCHER

Abstract. Establishing the presence of birds, determining their breeding status, and providing consistent, standardized data constitute the rationale behind accepted protocols for monitoring populations of *Empidonax traillii extimus*. During 1999 and 2000, we compared results from inexperienced flycatcher surveyors with those from an experienced surveyor on the same (well-studied) site along the Gila River in southwestern New Mexico. We compared these survey results with the best composite estimate based on all available data. Both experienced and inexperienced surveyors detected the species during both summers; however a pronounced difference in estimated numbers of breeding territories existed between the two groups in both years. The experienced surveyor detected pairs by the use of the soft *whitt* call, while the inexperienced persons relied on the *fitz-bew* call. Data obtained show that the field application of the prescribed protocol can lead to errors in estimating the actual numbers of existing territories.

Key Words: *Empidonax traillii extimus;* observer experience; Southwestern Willow Flycatcher; survey techniques; vocalizations.

Following the listing of the Southwestern Willow Flycatcher (*Empidonax traillii extimus*) as an endangered species by the U.S. Fish and Wildlife Service in 1995 (USFWS 1995), it became imperative to identify and monitor potential breeding areas for the occurrence of this species. Hence, a survey protocol was developed and implemented as a method to assist in the effective management and conservation of this species (Tibbitts et al. 1994, Sogge et al. 1997a). The primary stated objectives of this protocol are "to provide a standardized survey technique to detect southwestern willow flycatchers, and determine breeding status, and provide consistent and standardized data reporting" (Sogge et al. 1997a:17). The protocol consists of using tape playback of Willow Flycatcher songs (*fitz-bew*) to elicit responses from flycatchers. Surveys are conducted at least once in each of three survey periods (15–31 May, 1–21 June, and 22 June–17 July).

The current protocol was primarily designed to determine presence/absence, rather than an accurate count of flycatchers at any given site (Sogge et al. 1997a). However, an important secondary application of the procedure was to estimate population size, as indicated by several of the authors' statements (Sogge et al. 1997a: pp. 17, 23, 25): "Surveys conducted by qualified personnel in a consistent and standardized manner will enable continued monitoring of general population trends at or between sites . . . . Extra visits provide a greater confidence about presence or absence of flycatchers at a site, as well as help in estimating the number of breeding territories or pairs. . . . Given sufficient time, effort and observation, it is usually possible to approximate the number of territories and pairs." The use of survey results to estimate flycatcher numbers is also promoted by the official protocol data form, which instructs surveyors to estimate the number of pairs and territories detected for each survey period. The protocol thus becomes the primary source of population estimates for *E. t. extimus* for most locations. For example, in 2000, survey crews from the Arizona Game and Fish Department surveyed 197 sites although more intensive monitoring occurred at only 13 (Paradzick et al. 2001).

Prior to the present study, no field evaluation of the efficacy of this protocol existed. Braden and McKernan (1998) argued that the current protocol is inadequate even for detecting flycatcher presence. Their arguments were problematic for several reasons delineated by Sogge et al. (1999) and did not constitute a true field test.

The purpose of this study is not to criticize the current protocol but to evaluate with a field test its efficacy, and how closely surveyors are actually following the protocol. As additional information on a species is made available through scientific studies, protocols can and should be modified accordingly. Indeed the current protocol evolved from the first version by Tibbitts et al. (1994). Our objectives were to evaluate the ability of the protocol to detect flycatcher presence and to assess the effect of observer experience on the ability to estimate the number of breeding pairs.

METHODS

The study plot, approximately 11 ha known to be occupied by Southwestern Willow Flycatchers (hereafter "Willow Flycatchers"), was adjacent to the Gila River, approximately 50 km northwest of Silver City, Grant County, New Mexico, at an elevation of 1366

m. The site supported deciduous woods, extending in a mainly north-south direction along the east bank of the river, characterized by boxelder (*Acer negundo*), velvet ash (*Fraxinus velutina*), Goodding's willow (*Salix gooddingii*) and Fremont's cottonwood (*Populus fremontii*).

Each summer (1999 and 2000) we chose two different persons with no experience in surveying flycatchers to participate in the study. We also selected an additional participant (the same in both years) who had surveyed the species in the Gila Valley from 1995 to 1998. Each surveyor in this study had some previous exposure to bird identification either through a college course in ornithology or from extensive field training involving local birds. Prior to the study, all participants attended a Southwestern Willow Flycatcher survey training workshop, which consisted of presentations on the status and distribution, habitats, biology, cowbird impact, field identification, vocalizations, protocol and data sheets, and permitting issues relating to this species. At the conclusion of the classroom presentations, participants traveled to a site previously occupied by Willow Flycatchers to gain additional field experience. Participants in the workshop exhibited a wide range of experience, physical abilities, and field identification skills. At the workshop, no attempt was made to assess these skills.

To evaluate how accurately individuals with workshop training followed the protocol, each individual, before censusing flycatchers, was introduced to the boundaries of the study plot and instructed to "follow the protocol." Surveyors were given a field map and required to plot their route through the study area for each survey, noting the positions of all Willow Flycatchers detected and the manner of detection (sight or vocalization). Participants were not allowed to discuss among themselves or with others any aspect of the study for the duration of each breeding season. At the conclusion of the field season, surveyors were interviewed about the methods and procedures they used to survey flycatchers.

During both seasons, a field crew from the Rocky Mountain Research Station not associated with the actual surveys periodically visited the study area to locate breeding territories, conduct nest searches, and monitor nests of Willow Flycatchers. Criteria used to designate a breeding pair were presence of an active nest or fledglings, or observations of nest-building or copulation. We then compared the resulting data to those compiled by all surveyors, and a final composite estimate of breeding pairs was based on both experienced and inexperienced surveyor estimates, known nests, and breeding territories.

RESULTS

Little difference in survey effort was found between the inexperienced and experienced surveyors. During the three survey periods prescribed by the protocol, each survey required approximately 2.2 hours to complete. Survey results did not appear to be influenced by amount of time spent surveying.

During the 1999 breeding season, both groups identified up to two individuals by *fitz-bew* calls per survey (Table 1). In 2000, both groups detected more birds by *fitz-bew* calls per survey, with more detections made by the inexperienced surveyors (Table 1), suggesting that more Willow Flycatchers were present on the study site that year. Only one of the four inexperienced observers noted "soft" *whitt* calls in either year, whereas the experienced observer based most of his flycatcher detections on soft *whitts*.

TABLE 1. NUMBER OF SOUTHWESTERN WILLOW FLYCATCHERS DETECTED BY INEXPERIENCED (S1 AND S2) AND EXPERIENCED (S3) OBSERVERS VIA *FITZ-BEW* VERSUS SOFT *WHITT* CALLS

	1999[a]			2000[a]		
	S1	S2	S3	S1	S2	S3
Detected by *fitz-bew*						
Survey period 1	1	2	1	4	8	3
Survey period 2	2	0	1	6	7	4
Survey period 3	0	1	1	0	6	2
Detected by *soft whitt*						
Survey period 1	0	1	0	0	0	8
Survey period 2	0	2	10	0	0	20
Survey period 3	0	1	10	0	0	12

[a] Composite estimates of the number of breeding pairs within the patch = 12 pairs in 1999, and 20 pairs in 2000.

Data from nest searches confirmed a population increase from 11 breeding pairs in 1999 to 16 in 2000 (based on 63 and 165 person-hours of search effort, respectively). Estimated population numbers each year, based on composite estimates from all sources, are presented in Table 1. Compared to these estimates, inexperienced surveyors detected 8–40% of breeding pairs present, whereas the experienced observer recorded 83–100% of breeding pairs.

DISCUSSION

Results of this study indicate that both inexperienced and experienced surveyors detected flycatchers at the study site during both 1999 and 2000; however, surveyors differed in their estimates of flycatcher abundance. This in itself is not a surprising finding, given that observer experience and training are well known sources of variability in avian survey protocols (Ralph and Scott 1981). However, the current Southwestern Willow Flycatcher survey protocol was developed with the intent that it could be effectively used—after a structured training workshop—by relatively inexperienced surveyors (M. Sogge, pers. comm.). Therefore, it is important to determine the degree to which survey results are influenced by use of inexperienced surveyors, and to identify ways in which the protocol could be made more effective.

The first survey of a given season is intended

to coincide with the period of highest singing rates of newly arrived males (Sogge et al. 1997). Field work during this period is important in identifying potential Willow Flycatcher habitat, but may not indicate resident pairs, as not all breeding birds have yet arrived on territory, and locally singing birds may be either migrants or unpaired males. Impressions gained over several summers along the Gila River indicate that Willow Flycatchers do not respond consistently to tape playback of *fitz-bew* calls during the second and third survey periods when pairs are actively breeding. Therefore, the possibility exists that reliance on the standard protocol may overlook numerous flycatcher pairs.

Willow Flycatchers are known to give a variety of vocalizations in addition to the *fitz-bew* call (Sedgwick 2000). Owing to difficulty in separating these calls from similar vocalizations of other species, the protocol precludes their use in confirming the presence of Willow Flycatchers (Sogge et al. 1997a:24). A *whitt* note is well known in this species (Sedgwick 2000), but we believe the situation is more complex, with this species producing a "soft" *whitt* that differs in intensity from a "hard" *whitt*, a distinction not made in Sogge et al. (1997a). The hard *whitt* is remarkably similar to calls of Brown-crested (*Myiarchus tyrannulus*) and Ash-throated (*Myiarchus cinerascens*) Flycatchers, and is used as an alarm call by all three species. The soft *whitt*, unique in this location to Willow Flycatchers, seems to function as a contact call between members of a pair (see Barlow and McGillivray 1983, as cited in Sedgwick 2000; Rourke et al. 1999). The soft *whitt*, easily distinguished from the more forceful hard note with practice, is useful in detecting Willow Flycatcher pairs in the field. However, we noted a difference in the detection abilities of inexperienced and experienced surveyors during both summers (Table 1), the experienced individual relying more on detecting flycatchers through soft *whitt* notes.

CONCLUSIONS

The small sample sizes in this study preclude statistical analyses. We recommend that our experiment be replicated elsewhere, particularly in other habitats as habitat structure may influence detectability of Willow Flycatchers. The following conclusions should therefore be considered tentative.

The Willow Flycatcher survey protocol appears to be an effective method of detecting the presence of flycatchers by both experienced and inexperienced surveyors, although the potential exists for missing flycatchers when solely the *fitz-bew* call is relied upon. Based upon the interpretation of all data, however, the protocol is not effective in estimating the number of breeding pairs by inexperienced surveyors or those relying only on *fitz-bew* calls. Sogge et al. (1997a) strongly encourage additional follow-up visits to confirm Willow Flycatcher presence (by *fitz-bew*) when *whitt*s are heard in the field, as those authors do not consider the *whitt*s to be diagnostic for that species (Sogge et al. 1997a: 20, 24). Such follow-ups are not mandated in the protocol, however, and our results, based upon post-survey interviews, suggest that follow-up visits are not carried out in the field.

The experienced surveyor in 2000 had the benefit of site familiarity and knowledge of previous bird locations, both of which would increase the ability to detect Willow Flycatchers. We were not able in this study to quantify the effect of site familiarity on ability to locate birds. However, we did note differences, based upon post-survey interviews, in the methods used by experienced and inexperienced surveyors that might explain why inexperienced surveyors detected fewer flycatchers. The experienced surveyor played the taped *fitz-bew* call less while listening more for soft *whitt*s. The failure of inexperienced surveyors to detect Willow Flycatchers by soft *whitt*s is probably a function of the protocol's emphasis on the *fitz-bew* vocalization and a lack of recognition of soft *whitt*s by these individuals.

The soft *whitt* can be an important tool in estimating the number of breeding pairs, especially in the latter two seasonal survey periods. Observers unfamiliar with this vocalization appeared to consistently underestimate the number of flycatchers within the survey area. Incorporation of the soft *whitt* calls—and perhaps other distinctive vocalizations, such as the *wheeo* call (see Rourke et al. 1999)—into future versions of survey protocol and training sessions may increase the reliability of population estimates derived from future surveys.

ACKNOWLEDGMENTS

We thank S. and J. Runyan for access to their private property; C. Wolf, M. Riciardelli, J. Garcia, C. Galaway, D. Hawksworth, H. Walker, A. Favis, K. Brodhead, W. R. Trussell, H. Woodward, R. Hunt, and G. Sadoti for field assistance; Western New Mexico University, USDA Forest Service, and Phelps Dodge Corporation for generous financial assistance; and D. Parker who alerted us to the differences in Willow Flycatcher vocalizations. Comments by D. A. Zimmerman, B. Kus, and two anonymous reviewers improved this manuscript.